BRITISH MEDICAL BULLETIN

VOLUME 9

1953

Advisory Editor

SIR WILLIAM Mac ARTHUR

*

PHYSIOLOGY OF VISION

Scientific Editor . . . . . W. S. STILES

*

SOME ASPECTS OF ENZYME RESEARCH

Scientific Editor . . . FRANK DICKENS

*

VIRUSES IN MEDICINE

Scientific Editor . . . FORREST FULTON

MEDICAL DEPARTMENT · THE BRITISH COUNCIL · 65 DAVIES STREET · LONDON W.1

	PAGE
Introduction — *Sir John Parsons*	1
Some problems concerning the formation and circulation of the aqueous humour — *Hugh Davson*	5
Factors controlling ocular accommodation — *E. F. Fincham*	18
The nature of refractive errors — *Arnold Sorsby*	22
The interpretation of spectral sensitivity curves — *H. J. A. Dartnall*	24
Chromatic aberration and the polychromatic theory — *Hamilton Hartridge*	31
Defective colour vision — *W. D. Wright*	36
Visual properties studied by subjective measurements on the colour-adapted eye — *W. S. Stiles*	41
Stimulation of the retina with light fields of small size — *L. C. Thomson*	50
Colour vision in the peripheral retina — *R. A. Weale*	55
The absolute sensitivity of the eye and the variation of visual acuity with intensity — *M. H. Pirenne*	61
Electric records from the vertebrate optic nerve — *W. A. H. Rushton*	68
Physiology of vision : a bibliography of British books and periodicals — *Katharine Tansley*	75

Scientific Editor : *W. S. Stiles*

Physiology of Vision

Until comparatively lately the study of the physiology of the special senses has been carried out by sporadic enthusiasts and has received little encouragement from physiologists, who were chiefly engaged in teaching undergraduate medical students. Recognition of the practical importance of the elucidation of visual problems led the Medical Research Council at the instigation of its secretary, Sir Walter Fletcher, to institute a Physiology of Vision Committee thirty years ago. The brilliant work of Lythgoe, a Beit scholar working under the aegis of the MRC, resulted in his being appointed to a permanent post on the staff. The establishment of the Institute of Ophthalmology in London has led to a research department under the direction of Sir Stewart Duke-Elder, and to the transference of the work of the Physiology of Vision Committee to the Institute. Several research members of the Institute have contributed to this symposium.

The nutrition of the eye is peculiar in that its optical functions demand clear dioptric media, free from blood vessels, and curved surfaces kept at constant tension. The intra-ocular pressure is higher than that of any other lymph space in the body. The necessary energy is derived ultimately from the blood pressure; and when it was found by Leber that the intra-ocular pressure follows passively all large changes in the carotid artery pressure it was concluded that the aqueous was a mere filtrate from the blood plasma and not a true secretion. So simple an explanation seemed improbable, and it was found to be incompatible with some of the facts. Attention then became focused upon osmotic conditions at the blood-aqueous barrier, and the whole armoury of recondite thermodynamics was brought to bear upon the subject. Opinion veered to the opposite extreme, and the aqueous was thought to be a mere dialysate; but this theory too has been found to be incompatible with some of the facts.

A great incentive to investigation of the intra-ocular pressure has been its pathological increase in cases of glaucoma, which so often leads to blindness. In most of these cases there is evidence of blockage of the angle of the anterior chamber. It would seem probable that at any rate in these cases filtration plays some part.

No theory yet satisfactorily explains the formation and circulation of the aqueous fluid. The true explanation can only be arrived at by patient and laborious biochemical and biophysical researches such as have been carried out during the last three years by Duke-Elder, Davson, Ambache and other members of the Physiology of Vision Research Unit.

Helmholtz's *Physiological optics* has been for many years the unrivalled treatise on the eye as an optical instrument. The first accurate measurements of the curvatures of the corneal and lental surfaces were made by his ophthalmometer. He also determined the dimensions of other optical constants. The sensitivity of the retina to *x* rays has made possible exact measurements of the lengths of living eyes in situ, as described in the paper in this Bulletin by Sorsby.

Helmholtz advanced the first adequate theory of accommodation, which, in spite of Tscherning's theory, is still most generally accepted. Fincham's researches showed the importance of normal differences of thickness of the different parts of the lens capsule. He has also shown that the accommodation-convergence reflex depends upon the obliquity of incidence of the light rays and to some extent on chromatic aberration. The importance of the exact direction of incidence of rays on the cones of the fovea was shown some years ago by the experiments of Stiles and Crawford, and Fincham has shown that the Stiles-Crawford effect plays a part in the accommodation reflex.

The great range of sensitivity of the retina from the condition of light-adaptation to that of maximum dark-adaptation has long attracted attention, and many estimates of the maximum sensitivity have been made, the latest by Pirenne. The threshold seems to require only very few quanta of energy.

Retinal sensitivity is not only quantitative, as in the scotopic condition and as experienced by differences of brightness in the photopic condition. In the latter it is also qualitative, as shown by differences of colour perception. Moreover different parts of the retina—foveal, paramacular, and peripheral—show differences of brightness and colour sensitivity under constant conditions of adaptation.

Flicker photometry has enabled very accurate estimates of brightness to be made, even under heterochromatic conditions; and this method has greatly facilitated comparisons of the responses from different parts of the retina. The fundamental work of Ives and other physicists dealt essentially with central vision, and is characterized by minute accuracy of physical measurement. It has to be remembered that differences of sensation cannot be estimated quantitatively except on the basis of the somewhat doubtful Fechner elaboration of Weber's law. The measurements are of the stimuli, and not of the sensations directly. Hence the importance of maximum accuracy and a full statement of all the physical conditions of the experiments. Much previous work by psychologists and physiologists has been vitiated by the absence of such records, so that it has been impossible to repeat the experiments under identical conditions. The early work of Lythgoe, in collaboration with K. Tansley, on the peripheral retina was characterized by such meticulous accuracy that it satisfied the stringent demands of physicists. All of Lythgoe's work was indeed characterized by the same attention to detail, and his good example has been followed by L. C. Thomson, Dartnall, Weale and other members of the Physiology of Vision Research Unit.

The gathering of physiologists from all parts of the world to the International Congress of Physiology at Oxford in 1947 included many workers on the physiology of vision. The opportunity was seized of inviting them to a Conference on Colour Vision, held at Cambridge in the following week. Among those who contributed papers were the anatomists Le Gros Clark, Polyak, and Willmer, the physicists Wright, Stiles, MacAdam, Holmes, Judd, and Pitt, the biochemists Hecht, Wald, and Morton, and the physiologists Granit, Pirenne, and Hartridge.

It was in 1877 that Boll discovered visual purple. Kühne's researches from 1878 onwards settled the importance of this substance, better called rhodopsin, as the photochemical intermediary in the stimulation of

the retina under low conditions of illumination. This view was strongly confirmed by the fact that the absorption curve of rhodopsin is identical with the scotopic luminosity curve (with its maximum very near 500 mμ.). So thorough were Kühne's researches that little attention was paid to its constitution until Lythgoe and Wald, simultaneously and independently in England and the United States of America, attacked the problem with new and more refined biochemical techniques. The role of vitamin A in its formation was proved and has been confirmed by Morton. These new discoveries have been of fundamental importance.

There has been a vast amount of work on colour vision and its defects, largely owing to the importance of colorimetry in industry. Much was done by Guild and others at the National Physical Laboratory (Great Britain) and by Judd and others at the Bureau of Standards (USA). The trichromatic theory, which has always appealed most strongly to physicists, has been the foundation of this work. Wright has continued the fundamental researches of his predecessors Abney and Watson at the Imperial College of Science, London. He has redetermined with much greater accuracy the three so-called " sensation curves," and he has also done much work on colour-blindness. He gave a reasoned and judicial statement of the present status of the trichromatic theory at the Conference, and he has contributed a paper on defective colour vision to this symposium.

The work of Granit, of Sweden, on the neurophysiology of vision has been epoch-making. Elaboration of the methods of investigation of the electrical responses of the retina to light, initiated by Holmgren, Dewar and McKendrick and others, has culminated in the exquisitely delicate method of obtaining the responses from single fibres of the optic nerve by means of microelectrodes—the subject of a paper by Rushton in this symposium. The method has been made possible by the use of electrotonic amplifiers, first used for visual problems by Adrian. Granit's discovery of dominator and modulator curves is of great significance, but has perhaps raised as many new problems as those it has elucidated. It has always been an insuperable difficulty to correlate the photopic luminosity curve with the three sensation curves of the trichromatic theory. It may be that the dominator curve throws some light on the luminosity curve as being in some sense independent ; but the modulator curves are not easily correlated with the theory. Most important of all perhaps is the proof by Granit of interaction in the form of inhibition, summation, facilitation and so on in the retina. The retina is indeed unique in being an isolated piece of brain, connected with the rest of the brain by only a single nerve tract, the so-called optic nerve.

These researches thus show that investigation of the problems of the physiology of vision results not only in throwing light on vision but has a wider significance for more general problems. The same will prove true of researches on the intra-ocular pressure when the mystery of the production of the aqueous is solved. They will almost certainly have a bearing on the formation of the cerebrospinal fluid and of lymph production in general.

J. Herbert Parsons

Much of the work in arranging for this symposium on the physiology of vision devolved on Dr E. A. Carmichael of the National Hospital, Queen Square, London, to whom we are indebted for his help and advice throughout.

Notes on Contributors

DR H. J. A. DARTNALL has, since 1948, been studying the physical chemistry of the pigment substances which mediate vision. This interest dates from the period 1934–37 when, in association with Sir Charles Goodeve and the late Dr R. J. Lythgoe, he investigated the photochemistry of visual purple. From 1937 to 1940 Dr Dartnall worked in an industrial chemical laboratory, and during the Second World War he was engaged on research in offensive and defensive aspects of chemical warfare. In 1946, as Chemical Warfare Adviser to the Middle East and Mediterranean Command of the Royal Air Force, he disposed of allied and enemy stocks of gas weapons in Egypt and Italy. He was engaged on problems of rocket propulsion until 1948, when he joined the Vision Research Unit (now the Group for Research in the Physiology of Vision) of the Medical Research Council.

DR HUGH DAVSON is now working independently for the Medical Research Council in the Department of Physiology at University College, London, on researches into the permeability of natural membranes and the mode of formation of intra-ocular fluid. He was a Beit Memorial Fellow at University College, and a Fellow of the Rockefeller Foundation at the University of Pennsylvania. After working as Associate Professor of Physiology in Dalhousie University, Halifax, Nova Scotia, he returned to Great Britain and became Assistant Director of the Medical Research Council's Ophthalmological Research Unit, from which he resigned in 1951. His publications include *The physiology of the eye* (1949), *Textbook of general physiology* (1951) and *The permeability of natural membranes* (2nd ed., 1952).

MR E. F. FINCHAM is Director of the Department of Ophthalmic Optics at the Institute of Ophthalmology, London. He was from 1919 to 1950 lecturer in physiological optics and the principles of ophthalmic instruments at the Northampton Polytechnic in London. Mr Fincham served on the Council of the Optical Society from 1925 to 1930 and was Vice-President from 1930 to 1933. For his work on the application of photography to the study of the eye and particularly to the changes in the crystalline lens in accommodation, he was given the Fellowship of the Royal Photographic Society in 1945. He has designed a number of ophthalmic instruments, the chief of which is the coincidence optometer for the objective determination of the refraction of the eye. Mr Fincham is the author of papers on the mechanism of accommodation, and his monograph, *The*

NOTES ON CONTRIBUTORS

mechanism of accommodation, was published as a supplement to the *British Journal of Ophthalmology* in 1937.

PROFESSOR HAMILTON HARTRIDGE, at present engaged in part-time research work, was educated at King's College, Cambridge. During the First World War he was Experimental Officer at Kingsnorth Airship Station (Royal Air Force), and from 1919 to 1927 was Lecturer in Special Senses and Senior Demonstrator in Physiology at the University of Cambridge. For the next 20 years he was Professor of Physiology at St Bartholomew's Hospital Medical College, London, and was the first Director of the Vision Research Unit of the Medical Research Council, 1947–51. From 1934 to 1944 Professor Hartridge was a Governor of Harrow School, and since 1946 has been Gresham Professor of Physic at Gresham College, London. Three years later he was appointed Emeritus Professor of Physiology at the University of London. In 1926 he was made a Fellow of the Royal Society. His published work has covered many subjects, including the special senses, colour vision, haemoglobin and the blood, the reversion spectroscope, and the optics of the microscope; his book, *Recent advances in the physiology of vision* (1950) was reviewed in *British Medical Bulletin*, 1951, **8**, 94.

SIR JOHN PARSONS is Consulting Surgeon to the Royal London Ophthalmic Hospital and Consulting Ophthalmic Surgeon to University College Hospital, London; in 1904 he delivered the Arris and Gale Lecture to the Royal College of Surgeons, and in 1921 became a Fellow of the Royal Society. He has served on several committees set up to deal with problems of factory lighting, causes and prevention of blindness, and optical and ophthalmological problems in industry and in the armed forces. In 1921 he was President of the Illuminating Engineering Society and from 1925 to 1927 President of the Ophthalmological Society. Sir John Parsons is an Honorary Fellow and past President of the Royal Society of Medicine, and an Honorary Member of both the American Medical Association and of the Australian Ophthalmological Society. He is author of a number of books on the eye, of which *Colour vision* (1924) and *Introduction to the theory of perception* (1927) are classical works on their subjects.

DR M. H. PIRENNE is now Lecturer in Physiology at the University of Aberdeen, where his work concerns research into the physiology of vision with special reference to the quantum theory of light. He studied physical chemistry at the University of Liége and worked on molecular physics under Professor P. Debye in Liége and Berlin. Later he went to the USA with a travelling fellowship. His research on vision, in the laboratory of the late Professor S. Hecht in New York, led to the study of the implications of the quantum theory for visual excitation. During the Second World War, as a member of the Belgian Forces in the United Kingdom, he was seconded to work in Cambridge for the Military Personnel Research Committee of the Medical Research Council. After the war Dr Pirenne was appointed Imperial Chemical Industries Fellow of the University of London and resumed work on quanta and vision at the Royal Institution and at University College, London. He is author of *The diffraction of x-rays and electrons by free molecules* (1946), *Vision and the eye* (1948), and of various papers which have appeared in scientific periodicals.

DR W. A. H. RUSHTON was educated at Emanuel College, Cambridge, and at University College Hospital, London. Since 1935 he has been Lecturer in Physiology at the University of Cambridge, and, since 1938, Director of Medical Studies at Trinity College, Cambridge, of which he is a Fellow. He is a member of the Physiological Society and of the Society for Experimental Biology, and in 1948 became a Fellow of the Royal Society. Dr Rushton is at present engaged upon the estimation of the concentration of visual purple in the living eye by an objective method. In this method, blue-green light is shone into the eye and the fraction diffusely reflected back is measured by a photoelectric method; changes in this fraction appear to be due to changes in the amount of visual purple through which the emergent light has twice passed. He has published papers in the *Journal of Physiology*, in the *Proceedings of the Royal Society* and in other journals, and has contributed a chapter on nerve conduction to *Modern trends in neurology* (1951).

PROFESSOR ARNOLD SORSBY, now Research Professor in Ophthalmology, Royal College of Surgeons of England and the Royal Eye Hospital, London (the Hospital being part of the King's College Hospital Group) since 1943, was formerly visiting ophthalmologist to White Oak Hospital for Chronic Eye Disease in Children, at Swanley, Kent. His main interests lie in the genetic aspects of ophthalmology, and fundus diseases. At present Professor Sorsby is engaged on work on the genetic aspects of refractive errors, and on hereditary eye disease in animals. His publications include *A short history of ophthalmology* (2nd ed., 1948), *Genetics in ophthalmology* (1951) and *Systemic ophthalmology* (1951).

DR W. S. STILES has recently taken charge of the Photometric Section of the National Physical Laboratory, Teddington, Middlesex, where he has been concerned for some years with problems of physiological optics and with the performance of the eye under various lighting conditions. His published work includes papers in the *Proceedings of the Royal Society* and elsewhere on the nature and effects of glare, the directional sensitivity of the retina, and the spectral sensitivities of the rod and cone mechanisms of vision. During the Second World War he worked on methods of visual scanning and other problems involving vision. Dr Stiles has previously contributed to the Bulletin a paper entitled "Wartime problems of glare and dazzle" (*Brit. med. Bull.* 1947, **5**, 50).

DR KATHARINE TANSLEY was Lecturer in Physiology at University College, London, from 1940 to 1942, where in 1941 she won the Schafer Prize for Physiology. After this she worked for three years as a Hailey Stewart Research Fellow at the Strangeways Research Laboratory, Cambridge, and in 1945 she became Reader in the physiology of sense organs at the University of Birmingham. Since 1948 she has been Wernher Research Fellow (Medical Research Council) at the Institute of Ophthalmology, London, where she is engaged in research on neutron cataract, microanatomy of the retina and optic tracts, and inherited retinal degenerations in animals. Dr Tansley was editor of the sense organs section of *British Abstracts* from 1939 to 1949 and since then has been physiological member of the editorial board of *Ophthalmic Literature*. She is the author of several papers on visual purple, dark-adaptation, structure and development of the retina and on kindred subjects, which have appeared in the medical press, and has revised a section, "The eye", in Winton & Bayliss's *Human physiology* (3rd ed., 1948).

DR L. C. THOMSON is Director of the Medical Research Council's Group for Research in the Physiology of Vision at the Institute of Ophthalmology, London, and has, for some years, been working with the colorimeter designed by Professor W. D. Wright and built at the Imperial College of Science and Technology, London. With this instrument Dr Thomson has conducted subjective experiments which have shown the behaviour of the eye when stimulated with test patches of small size, and his article in this number of the Bulletin is a summary of this work. His main interest now is the electrophysiology of the eye and the optic nerve. Before joining the staff of the Medical Research Council in 1947, he was a Lecturer in Physiology at Guy's Hospital Medical School, London. He is the author of several chapters on the physiology of vision in *Colour in theory and practice*, edited by H. D. Murray (1952 edition), and also a paper entitled "A review of ideas on colour perception" (*Ophthal. Lit.* 1952, **6**, 3).

MR R. A. WEALE is at present Physicist to the Medical Research Council's Group for Research in the Physiology of Vision at the Institute of Ophthalmology, London, where his work concerns the investigation of total colour-blindness (cone monochromatism) and the effect on sensitivity of light absorption by the ocular media. He was formerly Research Assistant to the British Iron and Steel Research Association, and Lecturer at the South-West Essex Technical College in London. He has made studies of the reflectivity of the cat's tapetum and has contributed, to the *Journal of Physiology*, papers on colour vision in the retinal periphery.

PROFESSOR W. D. WRIGHT is at present Professor of Technical Optics at the Imperial College of Science and Technology, London. He graduated at Imperial College and in 1926 commenced research for the Medical Research Council on normal colour vision, the results of which were published in the Medical Research Council's Special Report No. 139, *Reports of the Committee on the Physiology of Vision : VII. A re-determination of the trichromatic mixture data* (1929, London). Subsequently he held research posts in industry in the USA and in Great Britain. He returned to the staff of Imperial College where, for the past 20 years, he has continued colour research, with financial support from the Medical Research Council and with the assistance and collaboration of postgraduate students. His work has been extended to include defective colour vision, colour adaptation, colour measurements, eye movements and stereoscopic vision. In 1951 he delivered the 17th Thomas Young Oration to the Physical Society, and in 1952 the Adolph Lomb Memorial Lecture to the Optical Society of America. Professor Wright's publications include *The measurement of colour* (1944), *Researches on normal and defective colour vision* (1946), and numerous papers in scientific journals.

SOME PROBLEMS CONCERNING THE FORMATION AND CIRCULATION OF THE AQUEOUS HUMOUR

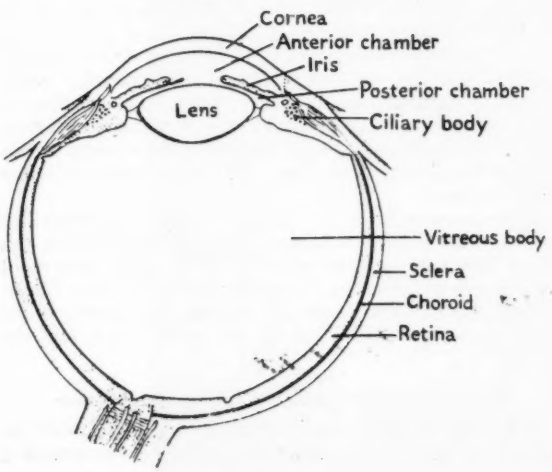
HUGH DAVSON D.Sc.

Medical Research Council
Department of Physiology
University College, London

1 Intra-ocular pressure
2 Nutrition of the lens and cornea
3 Chemical composition of the aqueous humour
4 Drainage and the blood-aqueous barrier
a Drainage
b The blood-aqueous barrier
5 Kinetics of the blood-aqueous barrier
6 Further analysis of kinetics of the blood-aqueous barrier
7 Effects of lipoid solubility
8 Some criticisms
References

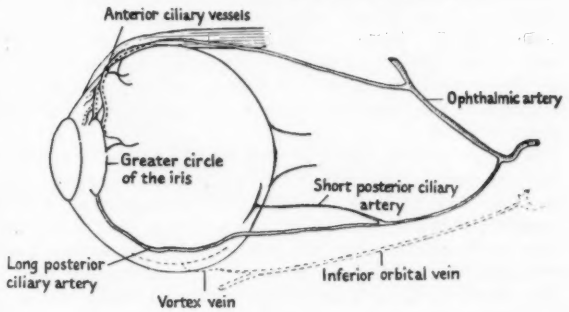
The aqueous humour is a clear colourless fluid occupying the anterior and posterior chambers of the eye (fig. 1). (By the posterior chamber is meant the space between the posterior aspect of the iris and the lens and zonule; the large space behind this and occupying, in man, the greatest part of the globe, is filled by a jelly, the vitreous body.) As an intermediary between the vascularized ciliary body and iris on the one hand, and the non-vascular lens and cornea on the other, the aqueous humour must fulfil an important role in the nutrition of these last-named structures. Moreover, the level of the intra-ocular pressure—an important clinical entity—is critically determined by the rates of production and drainage of this fluid, so that an exact knowledge of the factors controlling its production and drainage is vitally important for the understanding of a number of clinical conditions, e.g., glaucoma, in which a high intra-ocular tension may be an outstanding feature. It will assist the uninitiated reader to appreciate the problems with which this article is concerned, if a few anatomical points are first considered.

FIG. 1. DIAGRAMMATIC MERIDIONAL SECTION THROUGH HUMAN EYE



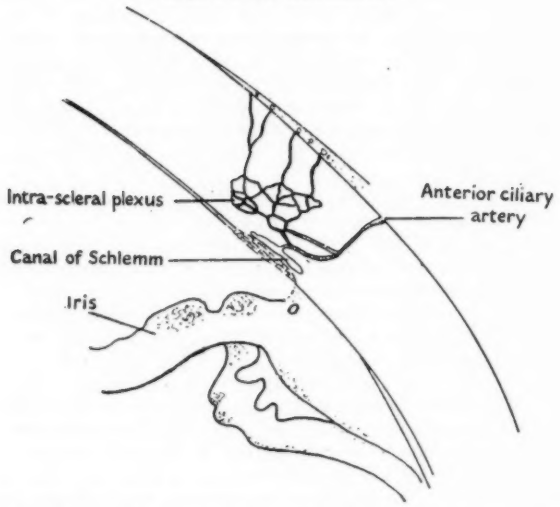
From fig. 1 it may be seen that the eye may be considered as made up of three coats: (i) the outermost—sclera and cornea, (ii) a middle vascular coat carrying the main blood supply to the eye and made up, from behind forward, of the choroid, ciliary body and iris, and (iii) an innermost layer—the highly complex retina in the more posterior region which is continued forward over the inner aspect of the ciliary body as a double layer of epithelium—the ciliary epithelium, and over the iris as the posterior iris endothelium. There is some question as to whether the anterior surface of the iris is also covered by an intact membrane, the anterior iris endothelium; if such a membrane is truly lacking, as claimed recently by Vrabec (1952), the iris represents a region in which the aqueous humour is in close functional contact with the capillaries of the vascular system—an important point when we consider the importance of secretory activity in its formation. The vascular supply with which we shall be primarily concerned consists of vessels supplying the ciliary body and the iris (fig. 2); these enter the eye as the long posterior, and the

FIG. 2. THE UVEAL CIRCULATION



anterior, ciliary arteries which anastomose at the root of the iris to form an arterial circle from which branches supply the various regions of the iris and ciliary body. The surface of the ciliary body is large, owing to the presence of numerous processes, each one of which is well supplied with blood by capillaries, and for this reason it is generally believed that it is this portion of the anterior uvea that is most directly concerned with the formation of the aqueous humour. The blood is drained away from the anterior uvea by the anterior ciliary veins—accompanying the anterior ciliary arteries—and the vortex veins—four large vessels which drain the great bulk of the blood from the eye. The anterior ciliary veins, although they drain only a small amount of the actual blood passing through the eye, are of great interest because they carry the blood from the intrascleral venous plexus, the vessels of which are in close relationship with the main drainage channel

FIG. 3. THE ANATOMICAL RELATIONSHIPS OF THE CANAL OF SCHLEMM



of the aqueous humour, the canal of Schlemm (fig. 3). This canal, seen in section in fig. 3, forms a circle (it was called the venous circle by Leber) at the junction between the cornea and sclera, and on its inner aspect comes into relationship with the aqueous humour through a spongy tissue—the scleral mesh-work. Presumably the aqueous humour percolates through the mesh-work and passes through the endothelial lining of the canal; from here it passes through collectors which empty into the vessels of the intrascleral venous plexus. Recently it has been shown (Ascher, 1942; Goldmann, 1946) that the aqueous humour may be drained to some extent by so-called " aqueous veins ", vessels that originate in the canal and connect with conjunctival vessels (Ashton, 1951).

Broadly speaking, then, we may take as our basic picture the continuous formation of a clear, almost protein-free, fluid by the ciliary body; this fluid, driven into the posterior chamber, causes the leakage of earlier-formed fluid from the angle of the anterior chamber, thereby giving rise to what has been described as the " circulation " of the intra-ocular fluid. This, of course, is a circulation only in a limited sense, the lost fluid being taken up by the blood and returning only after it has been freshly elaborated as a characteristic secretion.

With such a picture in mind we may briefly indicate some of the more obvious general problems that require solution and then pass to a detailed consideration of some special points with which the author and his colleagues (W. S. Duke-Elder, D. P. Greaves, A. Huber, M. Langham, P. A. Matchett, D. M. Maurice, J. R. E. Roberts, E. J. Ross, T. L. Thomassen and A. M. Woodin) have been specially concerned.

1. Intra-Ocular Pressure

The intra-ocular pressure will in large measure be determined by the forces driving aqueous humour into the eye and the ease with which it may escape. The latter factor, involving as it does purely mechanical considerations, is better understood than the former, which requires a knowledge of secretory mechanisms far in advance of the achievements of general physiology to date. Since the canal of Schlemm empties into the venous system of the eye, the pressure in the latter must play an important role in drainage and consequently exert an influence on the intra-ocular pressure. Such a view has been recently emphasized by Thomassen (1947) who has shown that some pathological variations in tension are actually *preceded* by a rise in venous pressure. Frictional resistance through the scleral mesh-work and the endothelium of the canal of Schlemm may be another factor, as emphasized by Goldmann (1948), who has brought a large amount of experimental support for a rise in resistance to outflow as a factor in glaucoma (see also Grant, 1951). The local blood pressure may influence the intra-ocular pressure in ways other than through drainage into the venous system; for example, a considerable capillary engorgement may produce a spectacular rise in intra-ocular pressure, attributable primarily to an exudation—normally held in check by the physiological tone of the vessels—of a protein-containing fluid into the anterior chamber from the iris; moreover, the increased volume of blood within the globe under these conditions must be a factor in the raising of the intra-ocular pressure. Alternatively, a constriction of the small vessels of the eye would lower the intra-ocular pressure, and it may well be that any nervous control of the intra-ocular pressure is exerted by way of vascular reflexes (Davson & Matchett, 1951; Davson & Huber, 1950; Greaves & Perkins, 1952).

2. Nutrition of the Lens and Cornea

The lens is a cellular structure and consequently requires a mechanism for the transport to it of glucose and other nutritional needs, such as amino-acids and vitamins; lactic acid and other products of metabolism must be removed. Whilst it is now clearly recognized that an adequate circulation of aqueous humour is essential for this purpose, the contribution of a defective circulation to the development of abnormalities in the lens seems not to have been thoroughly investigated; it is of interest in this connexion that, within an hour or so of the death of an animal, the lens already shows degenerative changes that are detectable by chemical means (Langham & Davson, 1949). The cornea is rather more favourably situated, in so far as it may derive nutrition both from the aqueous humour and from the vascular system at its periphery; removal of material from the aqueous humour will be retarded by the presence of the endothelial lining, whilst the obtaining of material from the capillaries of the limbal network vessels involves diffusion over long stretches of probably highly viscous corneal stroma. A recent and valuable contribution to our knowledge of the diffusion relationships between cornea, on the one hand, and aqueous humour and the limbal vessels on the other, is that of Maurice (1951), who made use of radioactive sodium.

These are one or two of the more general problems that present themselves to the student of the physiology of the eye; let us pass now to some more special questions that merit attention in so far as they reveal our ignorance of the fundamental mechanisms concerned in the formation of the aqueous humour and in exchanges between this and the blood.

3. Chemical Composition of the Aqueous Humour

We have stated, in our preliminary summary, that the aqueous humour is elaborated as a secretion; before accepting

MERIDIONAL SECTIONS OF FROZEN EYES

Hugh Davson

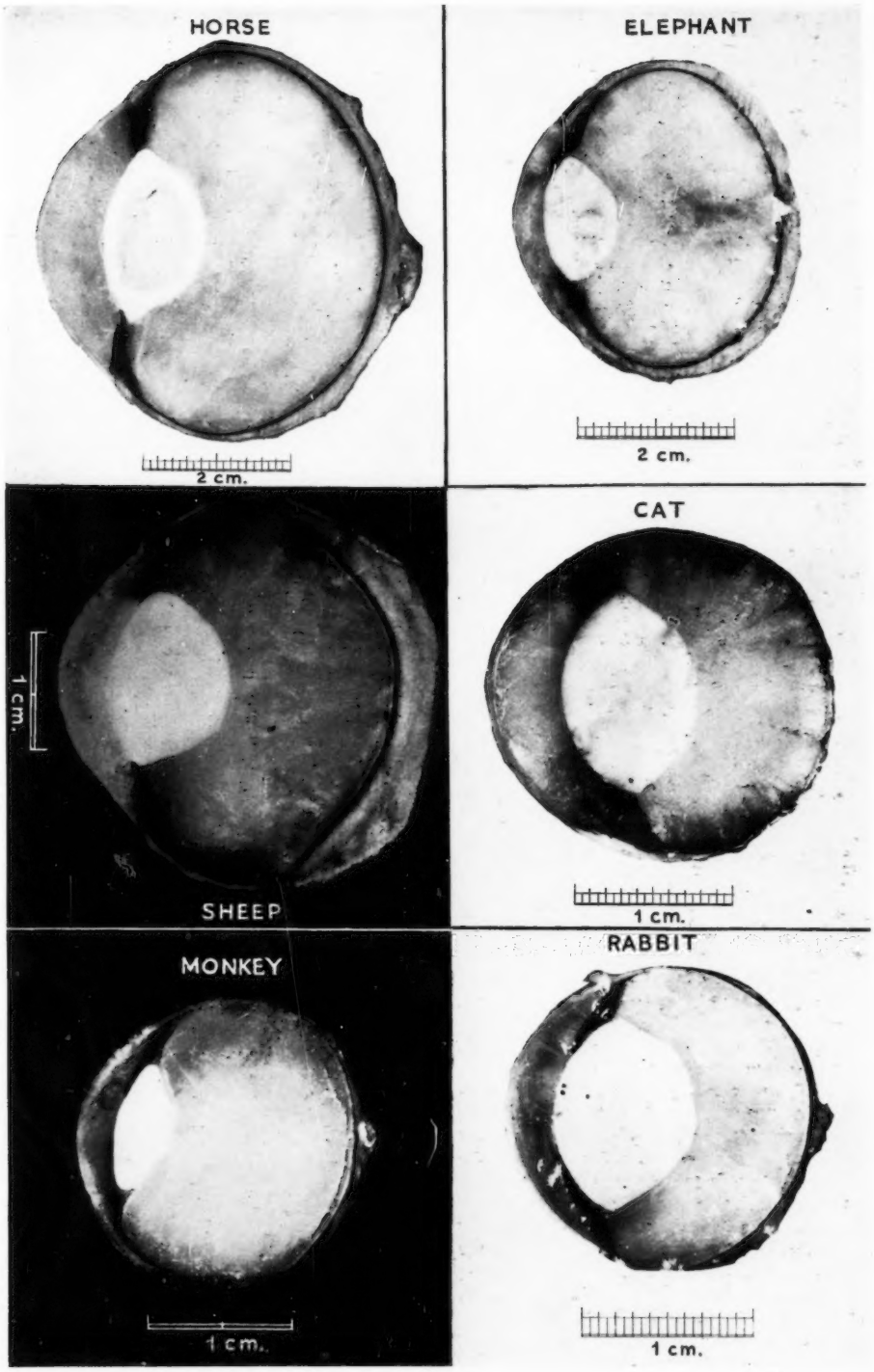
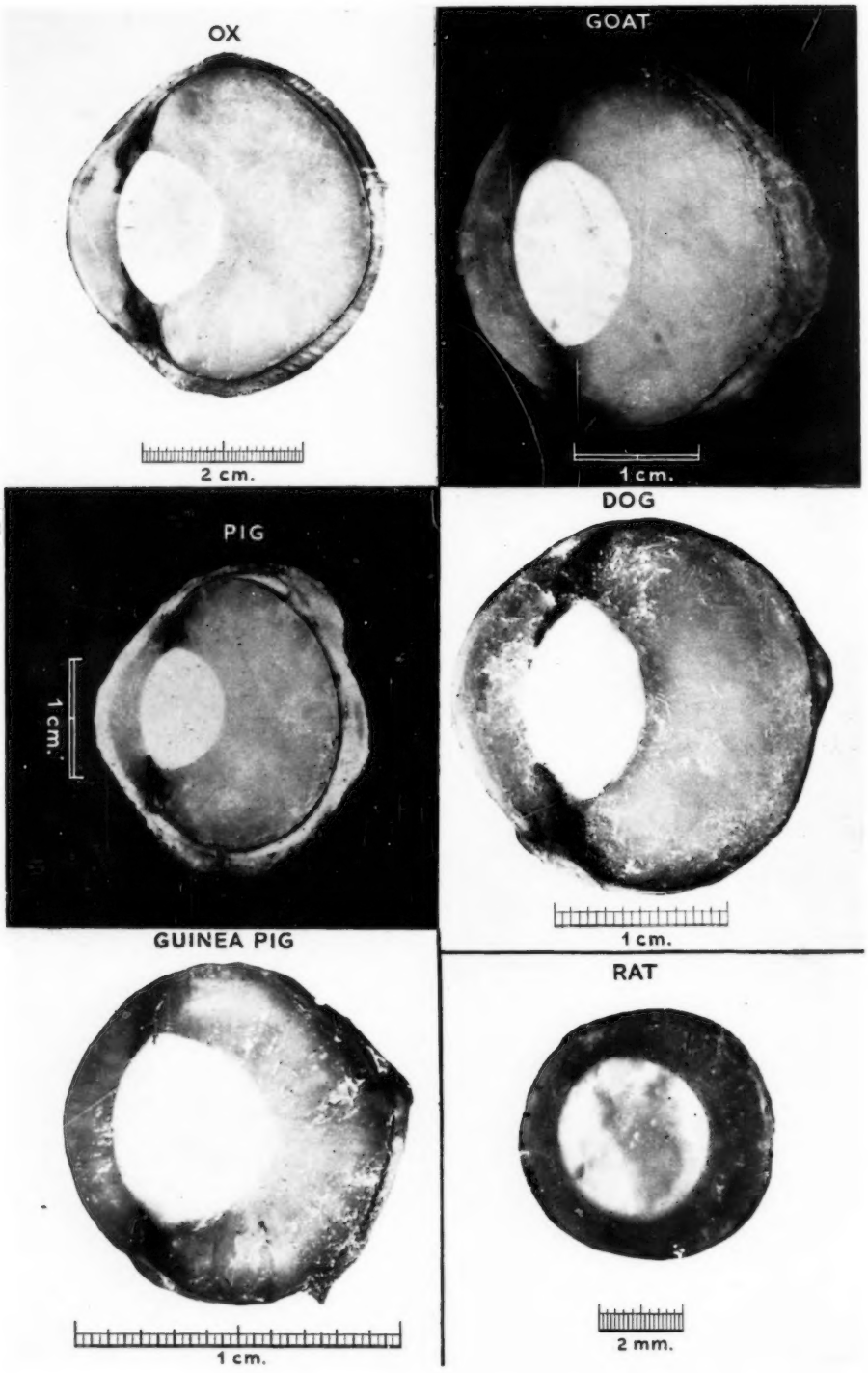


PLATE II

MERIDIONAL SECTIONS OF FROZEN EYES

Hugh Davson

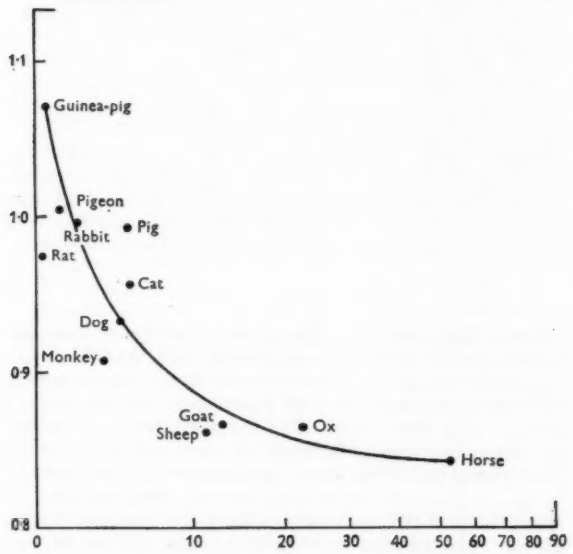


this statement, the critical inquirer will want to know whether simple physico-chemical factors are not adequate to explain its formation—that is, whether the fluid could be regarded as a filtrate from the blood plasma in the capillaries of the ciliary body and the iris. Such an hypothesis has indeed been put forward more than once, e.g., by Leber in 1895, but recent studies have shown that the chemical composition of the fluid diverges too strongly from that of an exudate from plasma to permit of its origin in this simple fashion. Particularly instructive in this respect are some studies of the distribution of sodium and chloride between plasma and aqueous humour: in the cat and dog it was shown that the distribution of these osmotically important ions differed appreciably from that pertaining between plasma and its dialysate (Hodgson, 1938; Davson, 1939; Davson & Weld, 1941); moreover, on dialyzing aqueous humour against plasma from the same animal, there was a movement of both Na^+ and Cl^- from the aqueous humour to the plasma, which indicated that the aqueous humour has a higher concentration of these ions than would be expected of a Donnan distribution (Davson, Duke-Elder & Maurice, 1949a). A remarkable feature of the ionic distribution is the very large deviation from equilibrium exhibited by the chloride ion; thus, in the dog, the ratio: $\frac{[Na]_{Pl}}{[Na]_{Aq}}$ was equal to 1.036, showing a discrepancy of 2.1% from the Donnan distribution of 1.057; the ratio: $\frac{[Cl]_{Pl}}{[Cl]_{Aq}}$ was equal to 0.91, giving a discrepancy of 7% from the Donnan distribution of 0.98. In the cat the discrepancy in the chloride ratio was smaller, 2.5%, whilst in the rabbit, as shown by Kinsey (1949) and later by Davson, Matchett & Roberts (1952), the chloride ratio is about 0.99, comparing with a Donnan distribution of 0.96 for this species. Thus, in this case there appears to be a deficiency of chloride in the aqueous humour. These remarkable specific differences tempted a further investigation on an extended series of species; the results are shown in fig. 4 (Davson *et al.* 1952), in which the chloride ratio has been plotted against the weight of the intra-ocular contents, and it will be seen that, in general, the larger the eye the smaller the ratio, i.e., the greater the excess of chloride in the aqueous humour. The guinea-pig eye, it will be observed, giving a ratio of 1.07, shows a marked *deficiency* of chloride in the aqueous humour.

It was thought, during the investigation of these distribution ratios, that an important factor in determining the species differences might be the varying proportion of lens to aqueous humour in the different eyes; thus, if the metabolism of the lens resulted in a modification of the ionic content of the aqueous humour, the greater the mass of lens in relation to that of the aqueous humour the greater the deviation from the true concentration of ions in the aqueous humour as originally elaborated. That the proportion of lens to aqueous humour may vary considerably from one species to another is evident from the photographs in Plates I & II, and it will be seen that, as the eye becomes smaller, the mass of the lens in relation to the whole eye becomes more significant. Nevertheless, studies on the rabbit, in which the lens of one eye was removed, failed to show a significant difference in chloride distribution between the two eyes—similar results had been found earlier in the dog (Davson & Weld, 1941). In such species as have been examined (Table I) the distribution of Na^+ showed no such marked variations; in general, dialysis of the aqueous humour against plasma resulted in a fall of the concentration in the aqueous humour of the order of 1%, suggesting that the accumulation of sodium in the aqueous humour is common to all species and that it occurs to about the same extent in them all. An interesting variation in the chloride distribution according to the age of the animal was observed: thus the ratio for a group of pups 3–5 months old was 0.97, whilst a group of adult dogs of the same breed gave a ratio of 0.93; a similar trend was observed in a comparison of calves with fully grown oxen. (Kinsey has shown that, in rabbits at birth, the blood-aqueous barrier is less specific than in the more mature animal (Kinsey & Jackson, 1949; Kinsey & Williamson, 1949).)

The full significance of these results will emerge only after

FIG. 4. VARIATION OF CHLORIDE DISTRIBUTION RATIO WITH SIZE OF THE EYE (Davson, Matchett & Roberts, 1952)



Abscissae: mass of intra-ocular contents—vitreous, aqueous and lens (arbitrary scale)

Ordinates: ratio of chloride in aqueous humour to chloride in plasma

TABLE I. RATIOS OF CONCENTRATIONS OF SODIUM IN THE WATER OF PLASMA AND OF AQUEOUS HUMOUR (Davson, Matchett & Roberts, unpublished)

Species	Number of animals	Ratio $\frac{[Na]_{Pl}}{[Na]_{Aq}}$
Horse	2	1.068
Steer	1	1.058
Sheep	8	1.06
Pig	3	1.07
Goat	6	1.064
Rabbit	19	1.061
Dog	18	1.036
Cat	9	1.042

further study along the same lines; but some of the problems they pose may be briefly indicated. First, we ask how a large excess of chloride, as, for example, in the horse, modifies the total osmotic pressure of the aqueous humour; this will depend on whether the excess is balanced by an excess of cations or by a deficiency of other anions, such as bicarbonate. Clearly, a first step is to carry out a comparative study of the distribution of bicarbonate and of the total osmotic pressure in the various species. A second point concerns the analogy between the cerebrospinal fluid (c.s.f.) and the aqueous humour; in their mode of origin and circulation, and in their general chemical composition, there are marked similarities. When the chloride distributions between c.s.f. and plasma are compared in different species, however, we find no such marked variation; thus in the rabbit, dog, ox and horse the ratios are: 0.825, 0.87, 0.85, and 0.84 respectively, comparing with ratios for the aqueous humour of: 0.99, 0.93, 0.86, and 0.84. These distributions indicate a fundamental point of difference between the two fluids, and it may be of significance that, as we progress from small to large animals, the distributions between aqueous humour and plasma, on the one hand, and c.s.f. and plasma on the other, approach each other.

The distribution of non-electrolytes raises some interesting questions; as a result of recent studies there seems little doubt that the concentrations of urea and glucose in the aqueous humour are smaller than those in the plasma. The deficiency of glucose can probably be completely explained by its utilization within the eye—e.g., by the lens; the deficiency of urea is harder to explain in this way, so that here we have more evidence for the intervention of secretory activity in the elaboration of the aqueous humour. In considering why a substance like urea should be selectively treated in this fashion, we must ask whether the low concentration of the aqueous humour is not an accident occasioned by the conditions of flow; such a consideration was put forward by Kinsey & Grant in 1942, and to appreciate the problem we must now pass to a consideration of the dynamics of formation of the fluid, and the nature of the blood-aqueous barrier.

4. Drainage and the Blood-Aqueous Barrier

a. Drainage

That the aqueous humour is continually drained away has been recognized on clinical grounds for many years, but a quantitative estimation of the rate at which it happens has been achieved only recently, although rough estimates—e.g., those of Troncoso (1921) and of Friedenwald & Pierce (1932)—have turned out to be remarkably near the truth. The problem of drainage is so closely related to that of the blood-aqueous barrier that the two must be considered together. When various substances are injected into the blood they penetrate into the aqueous humour at varying rates, ranging from the infinitely slow to the immensely rapid; this phenomenon has given rise to the term "blood-aqueous barrier", implying as it does that there is a hindrance to diffusion from the blood through the various structures that separate it from the aqueous humour. Thus inulin fails to pass from blood to aqueous humour at a measurable rate, whilst ethanol and water pass extremely rapidly. The first quantitative studies on this barrier were those of Davson & Quilliam (1940), who derived a simple equation to describe the process, containing a transfer constant, k, which measured the rate at which the substance came to equilibrium with the aqueous humour:

$$\frac{dC_{Aq}}{dt} = k(C_{Pl} - C_{Aq}) \qquad (1)$$

k being related to a true permeability constant, K, by the relationship: $k = \frac{KA}{V}$, where A is the area over which penetration takes place, and V the volume of the aqueous humour. C_{Pl} and C_{Aq} are the concentrations in the plasma and aqueous humour respectively.

In the derivation of this equation it was assumed that the factors favouring penetration were equal to the factors favouring the exit from the eye, as in a simple permeability process. Kinsey & Grant (1942) pointed out, however, that the continuous drainage away of the aqueous humour would favour the exit process more strongly than the inward penetration, so that an equation of the form:

$$\frac{dC_{Aq}}{dt} = k_{in}C_{Pl} - k_{out}C_{Aq} \qquad (2)$$

would be more appropriate. If k_{out} were greater than k_{in}, at infinite time we should have:

$$\frac{dC_{Aq}}{dt} = 0 = k_{in}C_{Pl} - k_{out}C_{Aq}, \text{ and hence } \frac{k_{in}}{k_{out}} = \frac{C_{Aq}}{C_{Pl}} = r \qquad (3)$$

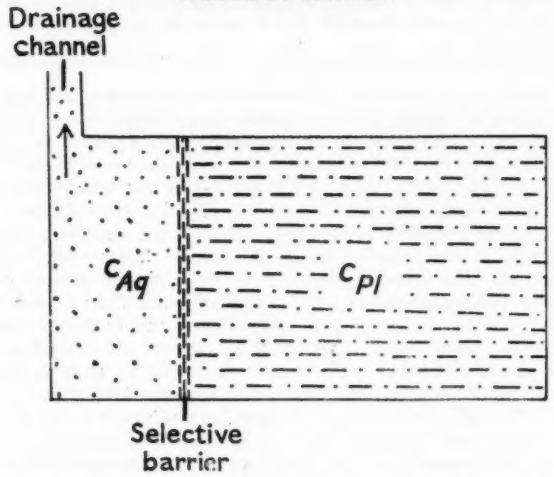
Under conditions in which the factors favouring entry and exit are equal in magnitude, r would be equal to unity, i.e., the penetrating substance would achieve equilibrium, as opposed to a steady state when r was less than unity. In this way it would be possible to explain the unequal distribution of urea between plasma and aqueous humour, provided of course that the energy necessary for maintaining this steady state were available. It has been shown that the deficiency of urea may be made up osmotically by the secretion of excess Na^+, so that the hypothesis is justified. We may now seek the physical basis for the condition that k_{out} should be greater than k_{in}. Drainage out of the eye is by flow across the endothelial lining of Schlemm's canal; if this endothelial lining has no very specific permeability characteristics—i.e., if most substances can cross it at the same rate with a low impedance to their flow, we will have here a significant factor favouring k_{out}. Thus we may imagine the system pictured in fig. 5: substances enter the fluid within the enclosed chamber across a highly specific barrier; continuously the fluid is withdrawn to be replaced by new fluid; the dynamics of the system would be described by the equation:

$$\frac{dC_{Aq}}{dt} = \left\{k_d(C_{Pl} - C_{Aq})\right\} - \left\{k_f C_{Aq} - nk_f C_{Pl}\right\} \qquad (4)$$

where the first term in brackets represents a simple permeability process, whilst the second represents the effects of the superimposed drainage and re-formation of the aqueous humour, k_f representing the fraction of the aqueous humour drained away in unit time. Thus, if all the constituents are carried through the endothelial lining of the canal of Schlemm at the same rate as water, the loss of concentration in unit time will be $k_f C_{Aq}$; if, on the other hand, the same does not apply to entry, the gain of concentration of a given substance due to the new fluid replacing that lost by drainage will be: $nk_f C_{Pl}$, where n is a fraction. Equation (4) has the same general form as that of equation (2); when n is equal to unity it acquires the still simpler form of equation (1).

Before discussing the mathematical treatment further, we

FIG. 5. SCHEMATIC DIAGRAM OF THE BLOOD-AQUEOUS BARRIER



C_{Aq}, C_{Pl}: concentrations of a substance in the aqueous humour and in the blood plasma, respectively

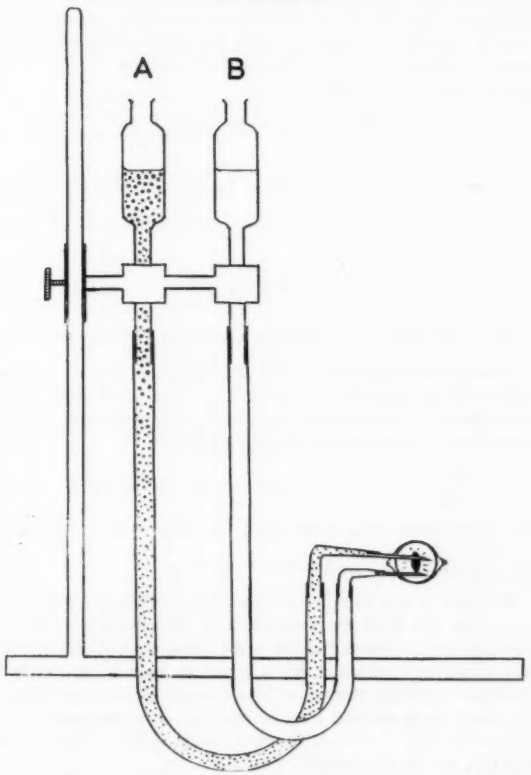
may proceed to the experimental verification of the primary hypothesis—namely, that drainage is a relatively unspecific process. This was carried out by the author some years ago by use of inulin, a polysaccharide of high molecular weight which had been shown earlier not to penetrate from blood to aqueous humour (Davson & Weld, 1941). This was introduced into the eye of a cat by inserting two needles, connected with reservoirs, into the anterior chamber as shown in fig. 6. Reservoir A contained a solution, approximating in composition to the aqueous humour, to which had been added inulin in a concentration of some 100 mg./100 ml.; reservoir B contained ordinary saline. The initial heights of the reservoirs were equal to the estimated intra-ocular pressure. By lowering B, a slow current of the inulin-containing saline could be run through the anterior chamber; this was allowed to proceed until the fluid within the chamber could be considered to have approximately the same inulin concentration as that in the ingoing fluid. The needles were then cut and sealed, and after about half an hour a third needle was introduced to withdraw the fluid in the anterior chamber. Analysis of the original inulin-containing fluid, and of that withdrawn, permitted a computation of the amount lost from the eye. Invariably it was found that the concentration of inulin had fallen considerably, in the region of 50% in half an hour. To ensure that the inulin was not simply diffusing back into the blood through the barrier that might have been broken down by the experimental treatment, a high concentration of inulin was maintained in the blood—greater than that in the anterior chamber—nevertheless, the concentration fell in the anterior chamber at a rate corresponding to a flow of some 1% per minute.

This experiment, whilst it answers unequivocally the question as to whether there is an unspecific flow out of the chamber, is unable to provide more than a very rough quantitative guide to the rate at which this flow takes place, since the eye could not be considered normal after the rather drastic treatment of inserting needles and flushing out the anterior chamber. Bárány & Kinsey (1949) devised a more physiological experiment; they injected *p*-amino-hippuric acid into the blood; this crosses the blood-aqueous barrier with difficulty, but within an hour there is an appreciable amount in the aqueous humour. Owing to its rapid elimination from the blood by the kidneys, the concentration in the plasma at the end of this period is very low, consequently the amount entering the eye after the first hour is very small. Owing to the drainage process, however, the *p*-amino-hippuric acid is being continually removed, and the rate of fall in its concentration gives an approximation to the rate of flow. Thus an animal is injected with, say, 1 g. of *p*-amino-hippuric acid and, after an hour, the aqueous humour of one eye is withdrawn; an hour later, the fluid from the other eye is taken. From the difference in concentration in the two fluids the rate of drainage may be approximately computed. The figure given by Bárány and Kinsey for the rabbit was in the region of 1% of the total volume of aqueous humour per minute, i.e., k_f was about 0.01 min.$^{-1}$.

b. The Blood-Aqueous Barrier

Having considered in some detail the relatively unspecific drainage process, we may now pass to the problem of

FIG. 6. EXPERIMENTAL STUDY OF DRAINAGE IN A CAT'S EYE



A: reservoir containing a solution approximating in composition to the aqueous humour, to which has been added inulin in a concentration of some 100 mg./100 ml.

B: reservoir containing ordinary saline

penetration of the blood-aqueous barrier. The permeability of this barrier, i.e., the relative difficulties various molecular species experience in traversing it, should provide some valuable information concerning the actual pathway from blood to aqueous humour—indeed, it was with this end in view that the author initiated a fairly comprehensive study of this aspect. The results of these studies (Davson *et al.* 1949b; Davson & Matchett, 1949; Ross, 1951) and of others from different laboratories, e.g., Palm (1947, 1948), Kinsey & Grant (1942) have brought out beyond reasonable doubt that the pathway—at any rate for the majority of substances—is definitely transcellular, i.e., that the selectivity of the barrier is so high that penetration through intercellular spaces must be ruled out as an important factor—at any rate where rapidly penetrating substances are concerned. If we accept the ciliary body as the main locus of formation of the aqueous humour, this would suggest that an important step in the passage of a molecule from the blood stream to the aqueous humour is constituted by the migration through the cells of the ciliary epithelium. If this view-point is correct, we may consider the kinetics of penetration in rather more detail than hitherto.

5. Kinetics of the Blood-Aqueous Barrier

Let us consider that a foreign substance is injected continuously into the blood to maintain a constant concentration, C_{Pl}, in the blood plasma. The equations hitherto used to describe penetration into the aqueous humour have the form of equation (2), namely:

$$\frac{dC_{Aq}}{dt} = k_{in}C_{Pl} - k_{out}C_{Aq} \qquad (2)$$

Since $r = \frac{C_{Aq}}{C_{Pl}}$ at infinite time, the integrated form of equation (2), C_{Pl} being constant and C_{Aq} being zero at zero time, is:

$$1 - \frac{C_{Aq}}{rC_{Pl}} = \text{exp.} -\frac{k_{in}}{r}t \qquad (5)$$

Thus the experimental determination of $\frac{C_{Aq}}{C_{Pl}}$ at a given time, t, together with its value at infinite time, permits the computation of the transfer constant, k_{in}. We have seen further that, when the penetration and exit by diffusion and flow are considered in more detail, the equation is:

$$\frac{dC_{Aq}}{dt} = k_d(C_{Pl} - C_{Aq}) - k_fC_{Aq} + nk_fC_{Pl} \qquad (4)$$

On integration this gives equation (5), with r equal to: $\frac{k_d + nk_f}{k_d + k_f}$, and k_{in} equal to $(k_d + nk_f)$.

We may note that r is less than unity when n is a fraction, i.e., when the fluid entering the eye to replace that lost by drainage has a concentration lower than that in the plasma. If n could be computed for a given substance, it would be possible to predict the value of r appropriate to any value of k_{in}, since k_f is known approximately. For example, if k_{in} were 0.02 min.$^{-1}$, and if k_f were taken as 0.01 min.$^{-1}$, and n as 0.5, we should have:

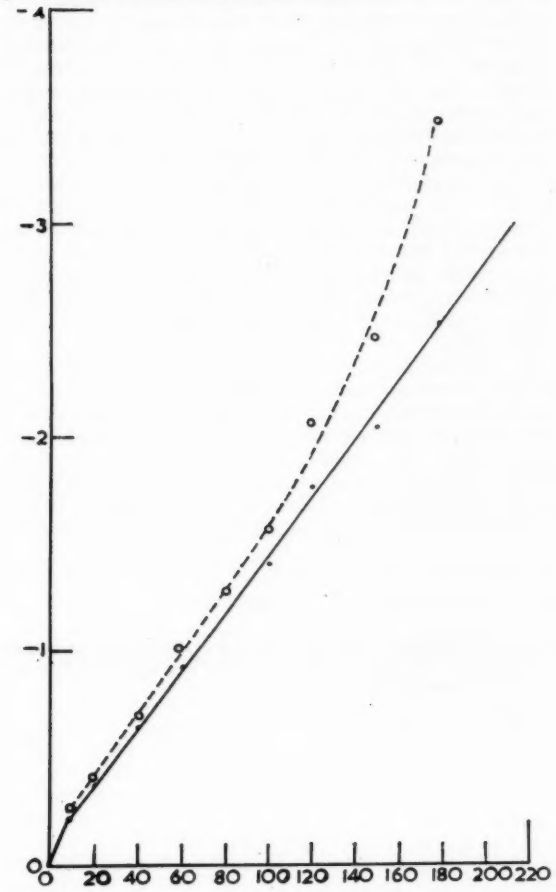
$$k_{in} = k_d + 0.5 \times 0.01 = 0.02. \quad \text{Whence } k_d = 0.015$$

$$\text{and } r = \frac{0.015 + 0.005}{0.015 + 0.01} = 0.8.$$

Unfortunately there is no valid a priori argument on the basis of which n may be estimated; we may note that Kinsey & Grant (1942) assumed that n was equal to zero, in which case r became equal to $\frac{k_d}{k_d + k_f}$. The theoretical objections to this treatment were put forward by the author, and later Kinsey & Bárány (1949) suggested that n should be the ratio of the transfer constants of the substance and of water, i.e., in the region of 1/20. On this assumption they constructed a curve relating k_{out} with r. The validity of this assumption is open to some objections on physico-chemical grounds, but it is worth putting it to an experimental test, especially as the matter is not purely of academic interest. Thus it has been argued that non-electrolytes enter by a simple process of diffusion because their values of r, computed from equation (4), are said to agree with their experimentally determined values; ions, on the other hand, are claimed not to conform to this relationship and therefore are presumed to enter the eye by a fundamentally different mechanism. Unfortunately, in the author's view, this argument has not been submitted to a sufficiently stringent test, and we may now proceed to a study of some non-electrolytes and weak electrolytes which, on a priori grounds, are unlikely to be subjected to specific secretory processes. The substances chosen were three sulphonamides and thiourea, all capable of being determined with ease and precision by standard chemical procedures. Sulphadiazine was chosen first, and since its rate of penetration into the eye, k_{in}, had been determined previously (Davson & Matchett, 1949)[1] it was sufficient to determine the final distribution ratio, r, at infinite time. To determine this with accuracy presents a number of problems that have not hitherto been taken seriously; thus the lens occupies a considerable volume in the anterior chamber of the rabbit's eye and it will clearly take a long time to come into equilibrium with the aqueous humour. Again, substances penetrate into the vitreous body more slowly than into the aqueous humour (Davson *et al.* 1949b), so that back-diffusion into the vitreous body will take place for some time, giving rise to a fictitious steady state lower than the true one (Maurice, 1951). The procedure adopted to overcome these difficulties was to maintain a high concentration in the blood for about two hours, to let it fall for one hour, and then to maintain it constant at the lower value for several further hours. To maintain constant levels in the blood a special injection machine was devised (Davson & Purvis, 1952). The results of these determinations gave a value of r varying between unity and 0.81, with an average value of 0.93. On the basis of the curve of Kinsey and Bárány, a value of 0.93 should correspond to a value of k_{in} equal to about 0.085 min.$^{-1}$; the value of k_{in} found by Davson & Matchett (1949) was approximately 0.016 min.$^{-1}$, i.e., it differs by a factor of over five from the one predicted. As we shall see, the value of 0.016 min.$^{-1}$, deduced from measurements over small time intervals, is too high, the true value being nearer to 0.01 min.$^{-1}$, so that the discrepancy is even greater. Sulphanilamide penetrates more rapidly than sulphadiazine; it is, moreover, almost completely undissociated at physiological pH, so that it cannot be considered as an electrolyte from a permeability aspect. According to Kinsey and Bárány it should give a value of r considerably below unity. Since sulphanilamide is a rather toxic

[1] That is, the transfer constant was computed on the assumption that r was equal to unity; if r is not equal to unity the computed constant is only approximately equal to k_{in}, but the error is very small.

FIG. 7. PENETRATION OF SULPHANILAMIDE INTO THE AQUEOUS HUMOUR OF THE RABBIT (Davson & Matchett, unpublished)



Abscissae: time (min.)

Ordinates: $\log_e\left(1 - \frac{C_{Aq}}{rC_{Pl}}\right)$

Continuous line: $r = 1.0$

Interrupted line: $r = 0.95$

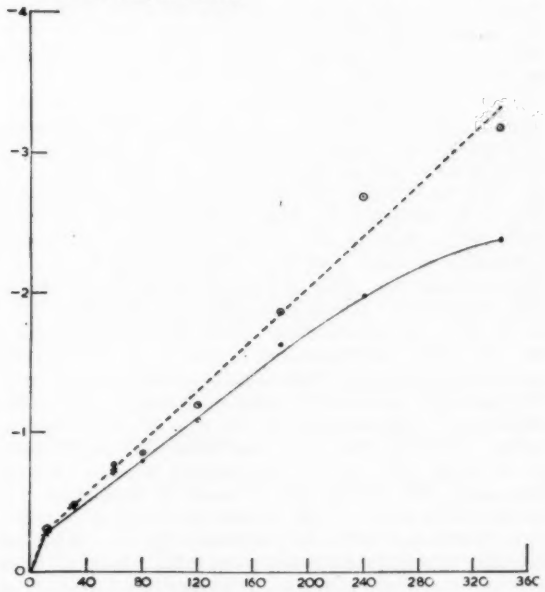
substance in very high concentrations, the technique employed with sulphadiazine—that is, the maintenance of a high concentration in order to raise the levels in the lens and vitreous body, followed by the maintenance of a rather lower one—was not followed. Instead, in groups of rabbits, the blood was maintained at a constant level, and the ratio $\frac{C_{Aq}}{C_{Pl}}$ was determined for different time intervals after the initial injection. From the trend of the points the final steady-state value of $\frac{C_{Aq}}{C_{Pl}}$ could be assessed with considerable accuracy. Thus the ratio should follow the equation:

$$\left(1 - \frac{C_{Aq}}{rC_{Pl}}\right) = \text{exp.}\frac{-k_{in}}{r}t \qquad (6)$$

so that plotting $\log_e\left(1 - \frac{C_{Aq}}{rC_{Pl}}\right)$ against time should give a straight line with slope equal to $\frac{-k_{in}}{r}$. By trial, a value of r should be found to give a straight-line relationship. Figure 7 gives the result of plotting the observed values when r is chosen as 0.95 and unity. Each point represents the mean of some five to six experiments. It is quite evident that a value of r equal to unity gives the better fit. In fig. 8 a similar treatment has been followed for thiourea; k_{in} in this case is considerably lower (0.009); the better fit is clearly with a value of r equal to 0.95.

The results quite unequivocally dispose of the theoretical treatment put forward by Kinsey and Bárány. Before developing an alternative treatment, more consistent with the facts, we must take note of an interesting feature of the graphs in figs. 7 and 8, and one that has perplexed the author for some time, that is, the considerable change in slope during the first ten to fifteen minutes. It had been considered earlier (Davson & Matchett, 1949) that this was largely an artifact resulting from the very high concentration of the sulphonamide in the arterial blood during the course of the initial injection, and also to the absence of adequate mixing of the fluid in the anterior chamber (Palm, 1947). The effects of the first factor have been eliminated by drawing blood continuously during and immediately after the initial injection—analysis of this sample gives an integrated value for the concentration during the first few minutes of the injection—and the points in

FIG. 8. PENETRATION OF THIOUREA INTO THE AQUEOUS HUMOUR OF THE RABBIT (Davson & Matchett, unpublished)



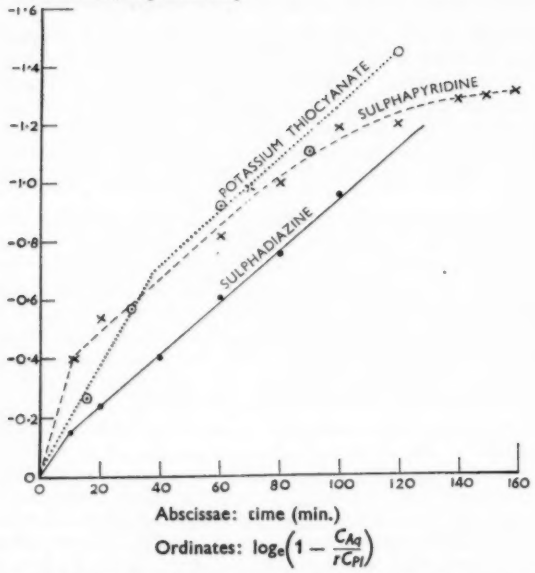
Abscissae: time (min.)

Ordinates: $\log_e\left(1 - \frac{C_{Aq}}{rC_{Pl}}\right)$

Continuous line: $r = 1.0$

Interrupted line: $r = 0.95$

FIG. 9. PENETRATION OF SULPHADIAZINE, SULPHAPYRIDINE AND THIOCYANATE INTO THE AQUEOUS HUMOUR OF THE RABBIT (Davson & Matchett, unpublished)

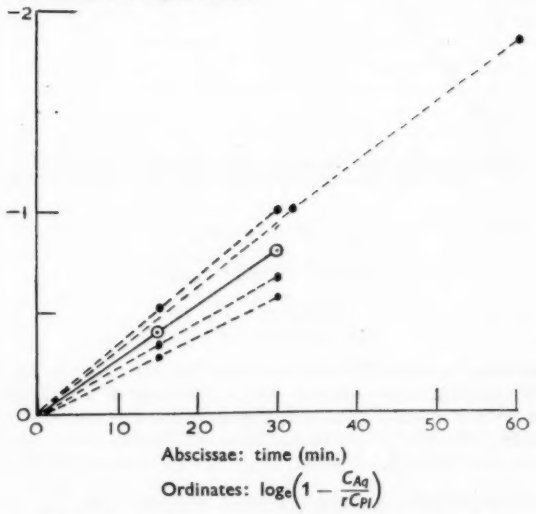


For sulphadiazine, r has been taken as 0.93; for the remaining substances, as unity.

figs. 7 and 8 have been corrected for this effect. That the effect is not peculiar to sulphanilamide and thiourea is shown by a study of several other compounds; thus curves for sulphadiazine, thiocyanate, and—in a much more marked degree—sulphapyridine, exhibit the same break (fig. 9). If the effect were due to inadequate mixing in the anterior chamber, we might expect it to be most marked with the more rapidly penetrating substances, yet, on comparing two homologous substances, thiourea and *n*-propyl-thiourea, we find that the effect is virtually absent in the more rapidly penetrating substance (fig. 10). Clearly we must look elsewhere for a cause of this break, and until a satisfactory explanation is found there is some uncertainty as to which slope measures the value of k_{in} appropriate to the steady-state distribution. Several possibilities occurred to us (Davson & Matchett, unpublished). First, there is the possibility that the injected substances were poisons to the secretory process, so that the rate of flow of aqueous humour was slowed. This possibility was tested by studying the effects of sulphanilamide and sulphapyridine on the rate of penetration of another substance, ^{24}Na, but it was found to be uninfluenced by the drugs. The results did emphasize, however, that the penetration of the isotope also exhibited the same break, and this became particularly evident when the results of an earlier study, by Maurice (1951), were replotted (fig. 11). A second possibility was that the substance was being converted in the body to a compound with a lower permeability constant; as a general explanation this must fail, since the isotope, ^{24}Na, showed the effect; nevertheless, the striking diminution in the rate of penetration of sulphapyridine, falling from an initial value of 0.04 min.$^{-1}$ to one of 0.0092 min.$^{-1}$ between 10 and 100 min. after the initial injection, suggested that, in this case at least, something of the sort was happening. Paper chromatography of the blood filtrate failed to reveal the presence of a derivative of sulphapyridine; but, by measuring the partition coefficient of free sulphapyridine between ether and plasma[2], definite evidence of the presence of a new, water-soluble, substance was obtained. Thus the partition coefficient between ether and plasma, when the sulphapyridine was added to this, was in one experiment 0.565; when the sulphapyridine had been injected into the animal, and the plasma removed after 100 min., the partition coefficient had fallen to about a quarter (0.122); in another the fall was from 0.52 to 0.22. The compound is probably a glucuronide of hydroxypyridine, giving the same chemical reaction in the Bratton and Marshall estimation, but, being much more water-soluble, penetrating the blood-aqueous barrier much more slowly, possibly not at all. Such a compound has, indeed, been isolated from the urine of dogs (Weber, Lalich & Major, 1943).

As indicated above, however, it is unlikely that this is a general phenomenon, but its observation in the case of sulphapyridine prompts caution in the interpretation of any experimentally observed value of r, the steady-state distribution ratio. Thus, without a thorough investigation of the kinetics of penetration of sulphapyridine, it might have been argued that substances penetrating very rapidly, i.e., with a k_{in} as high as 0.04 min.$^{-1}$, could have a value of r equal to 0.72, the steady-state ratio which $\frac{C_{Aq}}{C_{Pl}}$ appears to approach as a result of the formation of this compound.

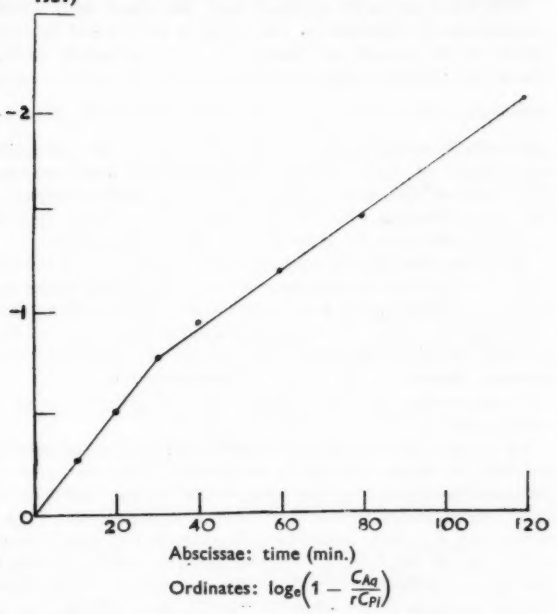
FIG. 10. PENETRATION OF *n*-PROPYL-THIOUREA INTO THE AQUEOUS HUMOUR OF THE RABBIT (Davson & Matchett, unpublished)



Continuous line: mean of four experiments; $r = 1.0$
Interrupted lines: individual experiments; $r = 1.0$

[2] Actually the partition coefficient between ether and a dialysate of the blood plasma was measured; this was because an appreciable fraction of the sulphapyridine is bound to protein and it was feared that the denaturation of the proteins consequent on shaking with ether would influence the amount bound and so give irregular results. It should be noted that in all these kinetic studies, it is only the unbound fraction of the substance that has been considered; this was determined by dialyzing plasma from the animal after injection with the appropriate substance.

FIG. 11. PENETRATION OF ^{24}Na INTO THE AQUEOUS HUMOUR OF THE RABBIT (replotted from Maurice, 1951)



What seems, at present, the most likely explanation of the break in the curves of penetration of so many substances is the presence of the small amount of fluid in the posterior chamber. This probably behaves as a "dead space", coming into a relatively rapid diffusion-equilibrium with the plasma in the ciliary processes and in the posterior layers of the iris. This would result in an initial rapid rise in the average concentration in the aqueous humour; if subsequent penetration were largely determined by the flow of aqueous humour through the pupil, the ensuing rise in C_{Aq} with time would follow a slower and more regular course. If we compare the penetration of two homologous substances, one penetrating relatively rapidly, e.g., *n*-propyl-thiourea with a k_{in} of 0.0267 min.$^{-1}$, and one penetrating more slowly, e.g., thiourea with a k_{in} of 0.009, we observe that *n*-propyl-thiourea barely shows a break in the plot of $\log_e\left(1-\frac{C_{Aq}}{rC_{Pl}}\right)$ against time, whereas it is quite marked with thiourea. The value of k_{in} for *n*-propyl-thiourea just mentioned was obtained from six experiments by Davson & Matchett (unpublished); it differs markedly from the figure given by Ross (1951), that is, 0.045. However, if one calculates a value from the curve presented in Ross's fig. 3, a lower figure of 0.036 is obtained, but still higher than the figure obtained in the author's laboratory. Unfortunately the substance is no longer obtainable so that further experiments are out of the question. We may anticipate the subsequent discussion and point out that the difference in k_{in} is due to a much more rapid diffusion across the iris, directly into the anterior chamber, in the case of *n*-propyl-thiourea. This rapid diffusion will mask the "dead space effect" of the posterior chamber, and thus explain the absence of an appreciable break in the graph. When, however, k_{in} becomes very high indeed, as with ethanol with a k_{in} of 0.07 min.$^{-1}$ (Palm, 1947), the defective mixing in the anterior chamber will cause a diminution in the rate of achievement of equilibrium, so that once again a break will occur in the graph of $\log_e\left(1-\frac{C_{Aq}}{rC_{Pl}}\right)$ against time; this, indeed, was observed by Palm.

Thus, in the case of thiourea, sulphadiazine, sulphanilamide, ^{24}Na, and a number of other substances penetrating with values of k_{in} in the region of 0.01–0.03 min.$^{-1}$, it will be the later slope of the graph that determines the value of r, the steady-state distribution ratio at infinite time, since, as the steady state is approached, the concentration in the aqueous humour approaches uniformity and the value in the posterior chamber. With very rapidly penetrating substances, such as ethanol and water (Kinsey, Grant & Cogan, 1942), the simple permeability equation—assuming as it does an adequate mixing of the contents of the anterior chamber—is invalid; since the value of r in these instances is certainly unity, it would be of academic interest to derive a more suitable equation to describe the kinetics of penetration.

6. Further Analysis of Kinetics of the Blood-Aqueous Barrier

We may now proceed to develop a more realistic treatment of penetration from blood into aqueous humour, taking into account the anatomical relationships of the posterior and anterior chambers, and see whether a theoretical relationship between r and k_{in} may be deduced which will give predictions that are not in such obvious contradiction to the facts as those of Kinsey and Bárány.

Let us consider that a foreign substance is injected continuously into the blood, as before. The substance, we may assume, diffuses into the cells of the secreting ciliary epithelium at a rate determined by the parameter in the equation:

$$\frac{dC_s}{dt} = k_s(C_{Pl} - C_s) \qquad (7)$$

where C_s is the concentration of the substance in the secretory cells at any time, t. The secreting cells, we may assume, eject continuously into the posterior chamber a fluid which we may call "primary aqueous humour", containing the substance in concentration, C_s; the rate of ejection will be equal to the rate of drainage, given by the parameter, k_f; it is the fraction of the total volume of aqueous humour drained away in one minute. Consequently the concentration within the secretory cells, at any time t, will be given by the differential equation:

$$\frac{dC_s}{dt} = k_s(C_{Pl} - C_s) - k_f C_s \qquad (8)$$

which gives on integration:

$$C_s = \frac{k_s C_{Pl}}{k_s + k_f}(1 - \text{exp.} - (k_s + k_f)t) \qquad (9)$$

According to this simple scheme—too simple, as we shall see—the appearance of the substance in the aqueous humour will be determined by the equation:

$$\frac{dC_{Aq}}{dt} = k_f C_s - k_f C_{Aq}$$

$$= \frac{k_s k_f}{k_s + k_f} C_{Pl}(1 - \text{exp.} - (k_s + k_f)t) - k_f C_{Aq} \qquad (10)$$

Since the flow constant, k_f, is presumably the same for all penetrating substances, differences in rate of appearance in the aqueous humour will be determined by differences in k_s; according to equation (10), the maximum rate of penetration into the aqueous humour will be given when $k_s = \infty$, when equation (10) becomes:

$$\frac{dC_{Aq}}{dt} = k_f C_{Pl} - k_f C_{Aq},$$

i.e., the transfer constant becomes equal to the flow constant. This would mean that the maximum transfer constant is limited to a value of 0.01 min.$^{-1}$, or thereabouts. In actual fact, substances with constants considerably larger than this have been found, e.g., ethanol; consequently equation (10) is inadequate as it stands. A more realistic equation would be given by postulating an additional permeability process—directly across the blood-aqueous barrier—in addition to that taking place from blood to secretory cells. Thus, if diffusion could take place from the iris directly into the anterior chamber, this would introduce a new factor allowing transfer constants of indefinitely high values. If we introduce, then, an additional parameter, k_d, to take care of this diffusion process, we have:

$$\frac{dC_{Aq}}{dt} = k_f C_s + k_d(C_{Pl} - C_{Aq}) - k_f C_{Aq} \qquad (11)$$

Eliminating C_s with the aid of equation (9) and integrating, we obtain:

$$\frac{C_{Aq}}{C_{Pl}} = \frac{Z+W}{X} - \frac{Z}{X-Y}e^{-Yt} + \left(\frac{Z}{X-Y} - \frac{Z+W}{X}\right)e^{-Xt} \qquad (12)$$

where $W = k_d$, $X = k_f + k_d$, $Y = k_s + k_f$ and $Z = \frac{k_f k_s}{k_f + k_s}$.

When $t = \infty$, $\frac{C_{Aq}}{C_{Pl}} = r = \frac{Z+W}{X}$.

If penetration into the aqueous humour followed equation (12) it would be impossible to characterize the rate of penetration by a single parameter, k_{in}, as has been done in the past when using the simpler equation (6); instead, three parameters would have to be used which cannot all be computed separately. Thus the usefulness of equation (12) is at present limited; nevertheless it is of value in indicating within what limits the simpler equation (6) is valid, and also as a means of assessing—tentatively at least—the relationship between the rate of inward penetration and the final distribution ratio, r.

We may note, first, that when k_s is very high, equation (12) simplifies to:

$$1 - \frac{C_{Aq}}{C_{Pl}} = \text{exp.} - (k_f + k_d)t,$$

i.e., it is identical with equation (6), the value of r being unity and k_{in} being equal to $(k_f + k_d)$. Thus the occurrence of substances penetrating in accordance with this simple equation, such as sulphanilamide, strongly suggests that, in their case, equilibration between plasma and secretory cells is very rapid indeed. This is understandable as the ciliary processes, which are lined by the secretory epithelium, are very richly supplied with blood, and the area-to-volume relationship of each individual cell is highly favourable to a rapid achievement of equilibrium. The diffusion process, represented by k_d, is assumed to occur over the iris, which is less highly vascularized, and has a much smaller area in relation to the volume of fluid with which it has to equilibrate.

This point is worth emphasizing; the speed with which equilibrium is attained in the systems considered here is given by the transfer constant; this is a composite entity, being made up of a permeability constant, K, and the area-to-volume relationship, $\frac{A}{V}$; two systems may have identical permeability constants but widely differing transfer constants as a result of their different area-to-volume relationships. In the case of diffusion from the iris into the aqueous humour, the area-to-volume relationship is likely to be highly unfavourable; the reverse will be true for the secretory cells.

We have seen that sulphanilamide, with a k_{in} of 0.0134 min.$^{-1}$, obeys the above simple relationship, whilst thiourea, with a k_{in} of 0.009, comes very close to obeying it, with a value of r equal to 0.95. Consequently the limiting value of k_{in} at which the simplified form of equation (12) is followed lies between these two values; let us suppose it is 0.012 min.$^{-1}$. This means that $k_f + k_d = 0.012$; if k_f is 0.01, this leaves a component for diffusion of 0.002 min.$^{-1}$.

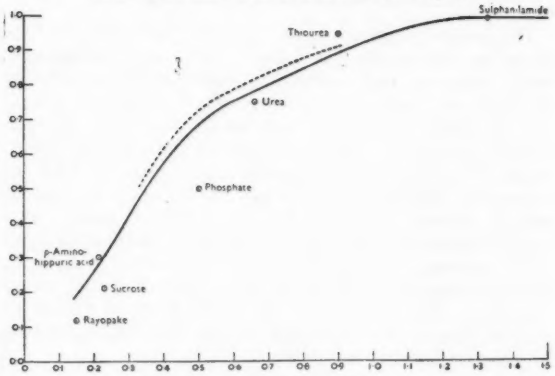
Let us now consider how less rapidly penetrating substances will behave, and try and predict a relationship between rate of penetration and r. As we have noted before, the rate of penetration cannot, in this case, be truly expressed as a single parameter, k_{in}, since equation (12) contains two exponential terms, one of which (exp. $-Yt$) becomes the more significant the more r differs from unity. Nevertheless it is sufficient for our purpose if we assess an "equivalent k_{in}", the parameter in the equation:

$$\frac{C_{Aq}}{C_{Pl}} = r\left(1 - \text{exp.} - \frac{k_{in}}{r}t\right)$$

obtained by substituting values of $\frac{C_{Aq}}{C_{Pl}}$ and t; for simplicity we may choose the value of t required for $\frac{C_{Aq}}{C_{Pl}}$ to be equal to $\frac{r}{2}$, i.e., the "half-life". Thus our problem is to assign values to k_s and k_d such that varying values of the half-life are obtained from equation (12); these half-lives are substituted in the above simple equation to give an "apparent k_{in}", i.e., the parameter that previous investigators would have assigned to a given substance if they had found this value of the half-life. The selection of these values of k_s and k_d need not be arbitrary; thus we know that when k_{in} is in the region of 0.012, k_s is infinity and k_d is 0.002; to obtain smaller values of k_{in}, and values of r less than unity, k_s must be given a series of decreasing finite values. If the diffusion factor, represented by k_d, consisted of penetration through water-filled intercellular pores, we could consider it to be tolerably constant by comparison with the variation of k_s, which involves penetration of cellular membranes; with this assumption we have merely to vary k_s from infinity to zero, and the result is shown as the continuous-line curve in fig. 12. On the curve have been put points to represent the experimentally determined values of "apparent k_{in}" and r, taken from the author's own work and from that of others.[3] With the exception of phosphate,

[3] I have used, with some misgivings, the values of k_{in} and r given in the literature for the slowly penetrating substances, sucrose, *p*-amino-hippuric acid, etc. In no case have the kinetics been adequately studied on the lines required by the considerations raised in this review, so that there is considerable uncertainty as to the exact values.

FIG. 12. VARIATION OF r WITH "APPARENT k_{in}" COMPUTED FROM EQUATION (12)



Abscissae: apparent $k_{in} \times 100$

Ordinates: $r = \left(\frac{C_{Aq}}{C_{Pl}}\right)$ at infinite time

Continuous line: computed curve assuming only k_s variable with k_d constant and equal to 0.002

Interrupted line: computed curve assuming k_s and k_d both variable with $k_d = 0.01\ k_s$

which lies right off the curve, the agreement between the experimental points and the theoretical curve is sufficiently good to suggest that the physical picture we have drawn—that is, the diffusion into secretory cells associated with a flow, superimposed on which is a diffusion process directly into the anterior chamber—may be the correct one for substances penetrating with values of "apparent k_{in}" varying from 0.002 min.$^{-1}$ to 0.012 min.$^{-1}$. The assumption that has not been completely justified is that the diffusion factor, represented by k_d, would remain constant over this range; if this factor were, indeed, due to diffusion through water-filled pores, we should have to introduce a correction for urea and thiourea, since their molecular weights are markedly different from those of the rest of the substances considered. Since k_d is necessarily small compared with k_s, at any rate over the range of values for "apparent k_{in}" considered here, these corrections will, however, not be very significant. Alternatively, we may assume that the diffusion factor is also determined by the penetration of cell membranes, e.g., of the iris endothelium; in this case the difference between k_s and k_d will be determined by the different area-to-volume relationships between secretory cells, on the one hand, and the surface of the iris and volume of aqueous humour, on the other. If this is true, k_d may be taken as some fraction of k_s; we may note that when k_s is equal to 0.2 min.$^{-1}$ and k_d is equal to 0.002, r is quite close to unity (0.96), i.e., that k_d, at the upper end of our curve, is about 1% of k_s. If, now, with this value of k_s as our starting-point, we decrease k_s and k_d in proportion, k_d being 1% of k_s, we get the curve in dotted lines (fig. 12); the difference between the curves is not great, certainly not great enough to warrant further discussion of the exact relationship between k_s and k_d for any given value of "apparent k_{in}".

Account must be taken of one further possibility: that the diffusion process is, in effect, a flow process, i.e., a continuous exudation from the iris capillaries into the anterior chamber; this is by no means a remote possibility, since it would seem, from a study of the muscle capillaries, that this is essentially the mode by which substances leave them. In this case equation (11) retains exactly the same form, with the experimentally determined k_f consisting of two components: one measuring the rate at which the secreting cells extrude fluid, k_f', the other, k_f'', the rate at which the filtration process from the iris takes place, and equal to k_d. Thus:

$$\frac{dC_{Aq}}{dt} = k_f'C_s + k_dC_{Pl} - k_f'C_{Aq} - k_f''C_{Aq}$$
$$= k_f'C_s + k_d(C_{Pl} - C_{Aq}) - k_f'C_{Aq}.$$

The important point to notice, however, is that the flow constant, k_f', is no longer the experimentally determined one ($k_f' + k_f''$), and now acquires a smaller value; this will have the effect of permitting values of r closer to unity with smaller values of k_{in} than those demanded by the original treatment.

The position with regard to sodium requires careful consideration. In general, we may state that the simple equations so far derived are fundamentally inapplicable, without modification, to the diffusion of ions, since this process is strongly affected by the existing potential gradients, which may be very large in the case of a secretory system such as the epithelium of the ciliary body. Nevertheless, in view of some recent discussions of the subject (e.g., Kinsey & Bárány, 1949), it is worth considering, on the evidence available, whether the behaviour of Na^+ differs strikingly from that of some non-electrolytes. The steady-state distribution ratio, r, has been found to be, by direct chemical analysis, 0.955 (Davson *et al.* 1952); this deviation from unity, however, is largely the result of the Gibbs-Donnan effect, which would impose a ratio of 0.945. Thus, in the absence of any secretory activity, there would be a potential equal to $\frac{RT}{F}\log_e 1.055$ directed so as to retard the penetration of Na^+, a potential which results in k_{in} being less than k_{out} by this factor. Consequently at electrochemical equilibrium a ratio, r, for Na^+ of 0.945 would correspond to a ratio of 1 for a non-electrolyte, entirely in virtue of the Donnan effect. To a first approximation, therefore, the curve of fig. 12, in order to be applicable to univalent positive ions, should be shifted downwards so that the limiting value of r is equal to 0.945; hence the observed k_{in}, if the same kinetics apply roughly, should be about 0.012 min.$^{-1}$. Maurice (1951) has examined the penetration of ^{24}Na more critically than anyone else, and his results have been presented earlier (fig. 11). If we ignore the early rapid rise, we note that the subsequent slope corresponds to a value of $\frac{k_{in}}{r}$ equal to 0.0148 min.$^{-1}$. In plotting this graph we have made use of a value of r equal to 0.90 instead of 0.955; the justification is as follows (Maurice, 1951): because, in the case of Na^+, penetration into the vitreous body is very slow, a definite proportion of the penetrating ^{24}Na is lost to the vitreous body in the early stages of penetration; in consequence, the concentration in the aqueous humour achieves, after four to five hours, a fictitious steady state with a value of r equal to about 0.9. (Maurice took a value of 0.8 for the fictitious steady-state ratio; however, on plotting his results as in fig. 11, it was found that a value of 0.9 gave a much better fit.) It is this fictitious steady-state value that should be used to determine the value of k_{in} rather than the true one, achieved only after some 20 hours. Consequently, k_{in} for Na^+ is 0.133 min.$^{-1}$, a value well within the limit demanded by fig. 12. We may conclude, therefore, that

there is nothing in the kinetics of penetration of Na^+ that permits a drastic distinction between it and such substances as thiourea, sulphanilamide, etc. This, of course, is not to deny that specific secretory processes are at work in transporting Na^+; we have already seen that its steady-state distribution ratio is actually higher than that demanded by the Donnan distribution; when it is realized, however, that this could be achieved by an active transport process that need only influence k_{in} by from 1% to 2%, it is clear that kinetic studies will have to be very much more refined than they are at present to permit of a differentiation between Na^+ and other substances to which such an active transport mechanism is not applicable.

7. Effects of Lipoid Solubility

We have seen that a value of r equal to unity will be given by any substance with a very high value of k_s, the transfer constant for penetration into the secreting cells, since in this case equation (12) takes the form:

$$\frac{C_{Aq}}{C_{Pl}} = \{1 - \exp. - (k_f + k_d)t\}.$$

We have seen also that this limiting condition is achieved with a value of k_{in}, i.e., $(k_f + k_d)$, equal to about 0.012 min.$^{-1}$. Let us consider, now, the penetration of substances with transfer constants higher than 0.012 min.$^{-1}$. Since k_f is taken as constant at 0.01 min.$^{-1}$, further increases of k_{in} must be the result of increases in k_d, the iris diffusion factor. Thus sulphanilamide has a value of k_{in} equal to 0.0134, giving a value of k_d equal to 0.0034; *n*-propyl-thiourea, with a k_{in} of 0.0267, gives a k_d of 0.0167; ethanol with a k_{in} of 0.07 (Palm, 1947) gives a k_d of 0.06. Such large increases in k_d, paralleling as they do the ether-water partition coefficients of the substances (0.19, 0.39, 1.9), must be a reflection of penetration through cellular membranes. Thus we may assume that, when the lipoid solubility is very low, penetration from the iris is by way of intercellular spaces only; as lipoid solubility increases, a point is reached where transcellular penetration becomes significant and finally, with a very high lipoid solubility, it dominates the picture. The influence of lipoid solubility has been discussed before (Davson & Matchett, 1949; Ross, 1951) but in this earlier work it was k_{in} that was compared with the ether-water partition coefficient, a procedure that led to an apparent anomaly and, as a result, to the present author's investigation of the possibilities of this new mathematical treatment. The anomaly is this: studies on plant and animal cells indicate that an increase of tenfold in the partition coefficient results in an increase in permeability constant of at least tenfold (vide, for example, Collander & Bärlund, 1933), whereas, in the case of the blood-aqueous barrier, although there is a fairly sound correlation between k_{in} and the partition coefficient, the factor by which k_{in} increases with a tenfold increase in the partition coefficient is much lower; in fact, about five. If, however, k_d is compared with the partition coefficient, the factor becomes much larger; in the case of the compounds cited, twenty. That it is k_d that should be considered in relation to the partition coefficient, rather than k_{in}, is clear from the mathematical formulation whereby k_{in} appears as $k_f + k_d$ when r is equal to unity; in physical terms we may say that with substances giving a value of r equal to, or near, unity, equilibration between plasma and secretory cells is so rapid that increasing the ease with which the solute penetrates the cells has no measurable influence on this rate; consequently the only way in which this increased permeability may manifest itself is in an increased value of k_d.

The present treatment of the kinetics of penetration of the blood-aqueous barrier has therefore three main advantages: firstly, it gives a rough quantitative prediction of the variation of r with the "apparent k_{in}" that is not grossly at variance with the facts; secondly, it clears up the anomaly in the relationship between k_{in} and the ether-water partition coefficient; thirdly, the postulated mechanisms are consistent with the anatomical relationships obtaining in the eye. It must be insisted, however, that no claim to finality is made; as a working hypothesis this treatment should be useful in planning further studies, for example, on the relationship of k_d to k_s, and, if it is sustained by further experiments, it should permit of a more exact study of the active transport of sodium or chloride, revealed by the studies of ionic distribution between the aqueous humour and plasma.

8. Some Criticisms

Some apparent objections may now be briefly indicated. Firstly, it may be argued that to assign a single parameter, k_d, to the direct process of permeation from plasma to aqueous humour is too great a simplification; instead, three processes should really be considered, to give three parameters as follows:

$$\frac{dC_{Aq}}{dt} = k_1(C_{Pl} - C'_{Aq}) + k_2(C_s - C''_{Aq}) + k_3(C_{Pl} - C''_{Aq}).$$

Here C_{Aq} is the average concentration in the aqueous humour when the fluids from anterior and posterior chambers are mixed; C'_{Aq} and C''_{Aq} are the concentrations in the anterior and posterior chambers respectively; the first term represents diffusion from the anterior surface of the iris, the second term represents a leak from the secreting cells, and the third term diffusion from the posterior surface of the iris. If our earlier arguments are correct, in the initial stages of penetration the second and third terms are indeed very significant, and lead to the rapid rise in the average value of C_{Aq} responsible for the break in the graph of $\log_e\left(1 - \frac{C_{Aq}}{rC_{Pl}}\right)$ against time (figs. 7–11); beyond this point, however, C''_{Aq} is probably fairly close to C_{Pl}, so that the third term becomes small. Moreover, since the volume of aqueous humour in the posterior chamber is so small, C''_{Aq} will rapidly approach C_s, the concentration of the fluid constantly extruded from the cells. This will tend to make the second term vanish. Thus, in the region beyond the break in the graph, the use of a single parameter to characterize the direct permeation from plasma to anterior aqueous humour is justified.

Secondly, we may point to the anomalous position occupied by the sugars; so far as the total reducing substances, mainly glucose, are concerned, r for the rabbit is 0.85 (Davson & Duke-Elder, 1948); k_{in} has been given as about 0.02 min.$^{-1}$ (Harris & Gehrsitz, 1949; Ross 1952a), so that, according to fig. 12, r should be equal to unity. However, this is to neglect the important influence of the lens and retina in utilizing glucose, and it seems very likely that the deviation of r from unity is entirely determined by this (Davson & Duke-Elder, 1948); thus the concentration of glucose in the vitreous body is nearly half that in the aqueous humour; the newly formed primary aqueous humour must lose a great deal as it flows in contact with this body; in flowing over the lens,

with a mass some two to three times the mass of the whole volume of aqueous humour, it must lose more, and thus a steady state in which C_{Aq} is considerably less than C_{Pl} is only to be expected. It is interesting in this connexion that, in the cat, k_{in} for glucose is about 0.012 min.$^{-1}$, i.e., nearly one-half of the figure for the rabbit; the steady-state ratio is lower (0.77) but not lower by so much as might be expected if the sole factor concerned were the relative values of k_{in}. The cat, however, has a much higher proportion of aqueous humour to lens than the rabbit (Plate I); moreover, the flow constants do not seem to differ markedly (Langham, 1951); consequently the lens will have a smaller effect on the concentration in the aqueous humour in the case of the cat than in that of the rabbit. It is worthy of note that the sugars are sharply differentiated from all the substances hitherto studied in both the cat and rabbit, in that the k_{in} for sugars is markedly different in the two species; thus ^{24}Na, urea, thiocyanate, sulphapyridine, and sulphathiazole give values of k_{in} that are roughly the same in the two species. It was suggested by Davson and Duke-Elder that the comparatively rapid rate of penetration of hexoses into the eye was the result of a membrane specialization, i.e., the development of active groups or carriers that would permit these highly lipoid-insoluble substances to cross the barrier with some ease; a similar type of membrane specialization had been observed in the cat erythrocyte (Davson & Reiner, 1942). If this is true, it is not surprising that hexoses have larger transfer constants in the rabbit's eye than in the cat's; the greater volume in proportion to the aqueous humour of the lens, in the case of the rabbit, demands a greater degree of membrane specialization to permit the adequate nutrition of the lens. Such a specialization would leave the transfer constants for other substances, such as sulphonamides, unchanged. A striking confirmation of this view-point has been provided by Ross (1952b), who showed that, in the alloxan-diabetic rabbit, k_{in} for glucose was very much reduced below normal; insulin produced a marked increase of k_{in}. With regard to hexoses other than glucose the position seems highly complex if some recent results of Harris & Gehrsitz (1951) are sustained by further experimentation; the extraordinary variations in r on passing from one sugar to another observed by these authors are not susceptible of an explanation on any simple kinetic theory (Davson, 1949), so that specific secretory mechanisms must be invoked.

Thirdly, the position of phosphate in relation to the curve of fig. 12 is difficult to explain; phosphate is, indeed, intimately connected with the metabolic processes occurring in the lens (Palm, 1948) and presumably, also, with the energy-giving reactions that make secretory activity possible in the ciliary body. At present we can say that the kinetics of penetration of phosphate would appear to be markedly different from those of the penetration of metabolically inert substances.

Finally, we must not ignore the possibility that a considerable portion of the resistance to penetration of a substance from blood to aqueous humour resides in the capillary endothelium of the ciliary body and the iris; such a possibility had occurred to the author some years ago and various attempts were made to establish the truth or otherwise of this proposition, but with inconclusive results. Recently Palm (1951) has brought forward some evidence in favour of it by studying the simultaneous concentrations of ^{24}Na in aqueous humour on the one hand and in the ciliary body and the iris on the other. There are some theoretical objections to the interpretation of the experimental results, so that we must postpone a decision until a critical experiment has been performed. We may note, however, that the same mathematical treatment would be approximately valid, k_s acquiring a new significance, representing the resistance to penetration of both the capillary endothelium and the secreting cells.

In conclusion, although this review has been very strictly confined to those problems with which the author is actively concerned, it reveals the existence of many gaps in our knowledge of those fundamental facts without which no theory of the secretory process can be elaborated with any hope of success.

ACKNOWLEDGEMENT

It is a pleasure to acknowledge the assistance of Miss P. A. Matchett in carrying out the hitherto unpublished work described in this article and in the preparation of the diagrams.

REFERENCES

Ascher, K. W. (1942) *Amer. J. Ophthal.* **25,** 31
Ashton, N. (1951) *Brit. J. Ophthal.* **35,** 291
Bárány, E. H. & Kinsey, V. E. (1949) *Amer. J. Ophthal.* **32,** 177
Collander, R. & Bärlund, H. (1933) *Acta bot. fenn.* **2,** 114
Davson, H. (1939) *J. Physiol.* **96,** 194
Davson, H. (1949) *Ophthal. Lit.* **3,** 254
Davson, H. & Duke-Elder, W. S. (1948) *J. Physiol.* **107,** 141
Davson, H., Duke-Elder, W. S. & Maurice, D. M. (1949a) *J. Physiol.* **109,** 32
Davson, H., Duke-Elder, W. S., Maurice, D. M., Ross, E. J. & Woodin, A. M. (1949b) *J. Physiol.* **108,** 203
Davson, H. & Huber, A. (1950) *Ophthalmologica,* **120,** 118
Davson, H. & Matchett, P. A. (1949) *J. Physiol.* **110,** 416
Davson, H. & Matchett, P. A. (1951) *J. Physiol.* **113,** 387
Davson, H., Matchett, P. A. & Roberts, J. R. E. (1952) *J. Physiol.* **116,** 47*P*
Davson, H. & Purvis, C. (1952) *J. Physiol.* **117,** 18*P*
Davson, H. & Quilliam, J. P. (1940) *J. Physiol.* **98,** 141
Davson, H. & Reiner, J. H. (1942) *J. cell. comp. Physiol.* **20,** 325
Davson, H. & Weld, C. B. (1941) *Amer. J. Physiol.* **134,** 1
Friedenwald, J. S. & Pierce, H. F. (1932) *Arch. Ophthal., Chicago,* **7,** 538 ; **8,** 9
Goldmann, H. (1946) *Ophthalmologica,* **111,** 146
Goldmann, H. (1948) *Ophthalmologica,* **116,** 195
Grant, W. M. (1951) *Arch. Ophthal., Chicago,* **46,** 113
Greaves, D. P. & Perkins, E. S. (1952) *Brit. J. Ophthal.* **36,** 258
Harris, J. E. & Gehrsitz, L. B. (1949) *Amer. J. Ophthal.* **32,** 167
Harris, J. E. & Gehrsitz, L. B. (1951) *Amer. J. Ophthal.* **34,** 113
Hodgson, T. H. (1938) *J. Physiol.* **94,** 118
Kinsey, V. E. (1949) *J. gen. Physiol.* **32,** 329
Kinsey, V. E. & Bárány, E. H. (1949) *Amer. J. Ophthal.* **32,** 189
Kinsey, V. E. & Grant, W. M. (1942) *J. gen. Physiol.* **26,** 119; 131
Kinsey, V. E., Grant, W. M. & Cogan, D. G. (1942) *Arch. Ophthal., Chicago,* **27,** 242
Kinsey, V. E. & Jackson, B. (1949) *Amer. J. Ophthal.* **32,** 374
Kinsey, V. E. & Williamson, M. B. (1949) *Amer. J. Ophthal.* **32,** 509
Langham, M. (1951) *Brit. J. Ophthal.* **35,** 409
Langham, M. & Davson, H. (1949) *Biochem. J.* **44,** 467
Leber, T. (1895) *v. Graefes Arch. Ophthal.* **41,** 235
Maurice, D. M. (1951) *J. Physiol.* **112,** 367
Palm, E. (1947) *Acta ophthal., Kbh.* **25,** 139
Palm, E. (1948) *Acta ophthal., Kbh.* Suppl. No. 32
Palm, E. (1951) *Acta ophthal., Kbh.* **29,** 269
Ross, E. J. (1951) *J. Physiol.* **112,** 229
Ross, E. J. (1952a) *Brit. J. Ophthal.* **36,** 41
Ross, E. J. (1952b) *J. Physiol.* **116,** 414
Thomassen, T. L. (1947) *Acta ophthal., Kbh.* **25,** 221
Troncoso, M. U. (1921) *Amer. J. Ophthal.* **4,** 321
Vrabec, F. (1952) *Ophthalmologica,* **123,** 20
Weber, C. J., Lalich, J. J. & Major, R. H. (1943) *Proc. Soc. exp. Biol., N.Y.* **53,** 190

FACTORS CONTROLLING OCULAR ACCOMMODATION

E. F. FINCHAM F.Inst.P. F.R.P.S.

Institute of Ophthalmology
London

1 Influence of ocular convergence on accommodation
2 Influence of light vergence on accommodation
3 Chromatic stimulus to accommodation
4 Achromatic stimulus to accommodation
5 Discussion
References

In changing from distant to near vision two adjustments of the visual apparatus are required. The first is convergence of the visual axes in order that the optical fields of the eyes may become sufficiently compatible to allow binocular fusion; the other, an increase in the dioptric power to focus the more divergent rays of light. The factors affecting the operation of the second of these adjustments, namely, accommodation, are the subject of this paper.

The innervations governing these adjustments in normal binocular vision arise mainly from the *consciousness* that the distance of the object of regard has changed. The most prominent factor among many combining to give this consciousness is the disparity of the retinal images. This disparity acting through the fusion centre stimulates the act of convergence. The extent to which these changes in convergence may themselves influence accommodation has been investigated and will be described here. Some earlier experiments on the accommodation-convergence relationship (Adamson & Fincham, 1939) showed that the link between these functions was sufficiently close for an increase or decrease in ocular convergence always to be accompanied by an increase or decrease in accommodation. Nevertheless it is known that this linkage is not inflexible. If in these experiments the divergence of the light is kept constant and the convergence of the eyes artificially stimulated, the accommodation is not so great as that accompanying such a degree of convergence under natural conditions. It appears, therefore, that while convergence causes accommodation to be increased, the constant " vergence " of the light in this instance tends to keep it unaltered. The inference is that not only does convergence of the eyes influence accommodation, but also that the degree of convergence or divergence of the light at the retina acts as a stimulus to the mechanism controlling accommodation. This action of light vergence (degree of convergence or divergence) as an independent stimulus to accommodation will also be described.

1. Influence of Ocular Convergence on Accommodation

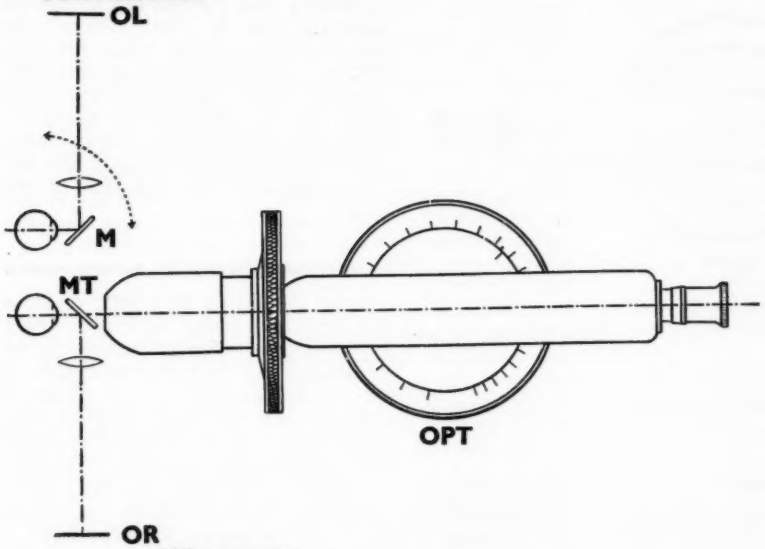
In the investigations to be described, the observations and measurements were made by an objective method. It is known that during accommodation, and particularly under experimental conditions in which accommodation and convergence are not equal, there can be an appreciable perceptual tolerance to departures from accurate focus of the retinal image. For this reason the method of determining the state of accommodation by the subjective judgement of definition is not dependable. Observations were made, therefore, with the coincidence optometer (Fincham, 1937).

The apparatus used in the subjective method of studying relationships between accommodation and convergence by earlier workers consisted of a form of Wheatstone's stereoscope: the haploscope of Hering. In this system each eye views an object through a lens and by reflection in a mirror at 45° to the visual axis. Each object, mirror and lens, is mounted upon a limb which may be rotated about the centre of rotation of the eye, and the objects may be moved along the axes of their corresponding lenses. Thus convergence and accommodation may be stimulated independently. The arrangement of the apparatus for the objective measurement of the effects is shown in fig. 1. It consists essentially of a combination of the haploscope and the objective optometer. One limb only of the haploscope is rotated, that before the eye which is not being measured; as convergence is a binocular act it is necessary to stimulate the rotation of one eye only, the other being kept aligned on the optometer.

To ensure that in this test the vergence of the light from the object produces no stimulus to accommodation, provision is made that the light shall enter the eyes through apertures of 0.5 mm. diameter, so that the effect of pin-hole vision is produced. The retinal images, therefore, are always well defined, except for some diffraction effect. This condition of the small entrance pupil is achieved by limiting the effective source of light illuminating the fixation object to a small area and arranging that this small source is imaged by the lens in the plane of the subject's pupil, while the fixation object is approximately conjugate with the retina.

The method of experiment is to set the haploscope so that, when the observer obtains binocular fusion of the objects, a definite amount of convergence is in play. Under these conditions the state of accommodation is measured with the optometer. These measurements show that in subjects in whom the extrinsic eye muscles are in a state of balance the act of convergence causes accommodation to be induced by an amount consistent with the convergence. For example, when, in order to obtain fusion, the eyes have to converge to a point at 1 m. distant, they accommodate 1 D.; convergence to 50 cm. is accompanied by 2 D. of accommodation. As convergence is made to increase, the induced accommodation increases in unison, up to an amount approaching the reciprocal of the distance of the near point. Here, owing probably to the unresponsiveness of the crystalline lens, additional convergence is not accompanied by any further accommodation. In cases where some physiological imbalance of the muscles exists, the equality between accommodation and convergence does not hold. In fig. 2 are shown the records of an exophoric and an esophoric subject. These results show that, because a greater innervation than normal is required to produce a given amount of convergence in cases of exophoria, the accompanying innervation to accommodation is also greater, which causes a relatively high response in that function. In esophoria, on the other hand, because the eyes have always a tendency to converge, the innervation required is less than normal, and so a weaker innervation is sent to the accommodation centre. The result in this case, as shown by the graph, is that convergence is in excess of accommodation.

FIG. 1. DIAGRAM (PLAN VIEW) OF APPARATUS FOR MEASURING CHANGES IN ACCOMMODATION INDUCED BY CHANGING CONVERGENCE



OPT: optometer
OR and OL: objects for right and left eye respectively
M: opaque mirror
MT: transparent mirror

2. Influence of Light Vergence on Accommodation

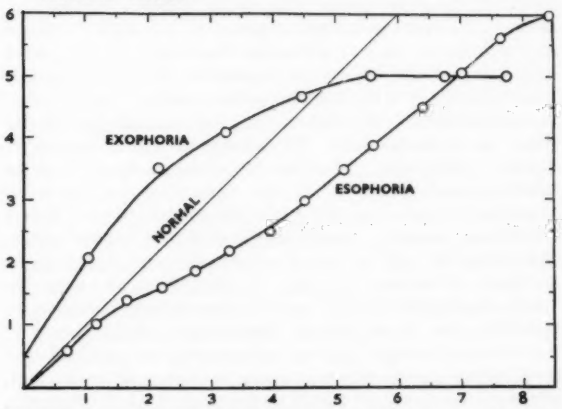
There is much subjective evidence that, in addition to the reaction to convergence, there is also a reflex reaction of the accommodative mechanism to alterations in the divergence of the rays of light from the object, without respect to the sense of distance. It is well known that, if in monocular vision a negative lens of low power is placed before the eye of a young normal subject, vision for distant objects is equally clear, and when the lens is removed clear vision is again obtained within a fraction of a second. The inference is that, when the lens causes the rays of light from the object to be divergent at the eye instead of parallel, the accommodation comes into action and adds sufficient power to the eye to focus the light sharply upon the retina. When the lens is then removed from before the accommodating eye the light is focused in front of the retina and so accommodation is relaxed and the normal condition restored. It is known also that young subjects with low degrees of hypermetropia constantly accommodate to correct the error until positive lenses are placed before the eyes, when the accommodation is automatically relaxed. In these cases there is no sense of change of distance and no conscious effort to maintain clear vision.

These observations lead to the conclusion that the visual mechanism must be sensitive to the vergence of the light striking the retina. If, alternatively, the mere loss of definition of the retinal image acted as a stimulus to the adjusting mechanism, then adjustment would be reached by a trial-and-error method, in the same way as we focus an optical instrument. This alternative is unlikely because, if vision is blurred by means of a positive lens causing the light to focus in front of the retina, an increase in accommodation would make the condition worse. We should thus expect fluctuation between accommodation and relaxation in an unsuccessful effort to obtain clear vision. However, it is generally agreed that positive lenses cause a constant relaxation of accommodation.

A true estimate of the behaviour of the mechanism can be obtained only by an objective method of observation and for this the coincidence optometer has been used. A reflected image of a small distant object was presented to the eye under test by means of a transparent mirror through which the optometer was directed (fig. 3). The divergence of the rays of light from the object was varied either by the interposition, between object and mirror, of lenses placed as near to the eye as possible or by the adjustment of a telescope of unit magnification placed in the same position.

Observation with the optometer confirmed that, when the rays of light from the object were made divergent at the eye, accommodation occurred. When the light was made parallel, accommodation relaxed. The presentation of convergent light to the normal eye produced no reaction. Various experiments on these lines (Fincham, 1951) showed that the accommodative mechanism was always adjusted in the correct direction; there was no doubt that the retina sent signals which informed the motor centre controlling accommodation whether it was necessary to increase or decrease the optical power of the eye. The supposition that the retina can detect and signal whether the rays of light reaching it are convergent or divergent is of course impossible, but the retina must detect and signal some characteristics of the image which convey this information.

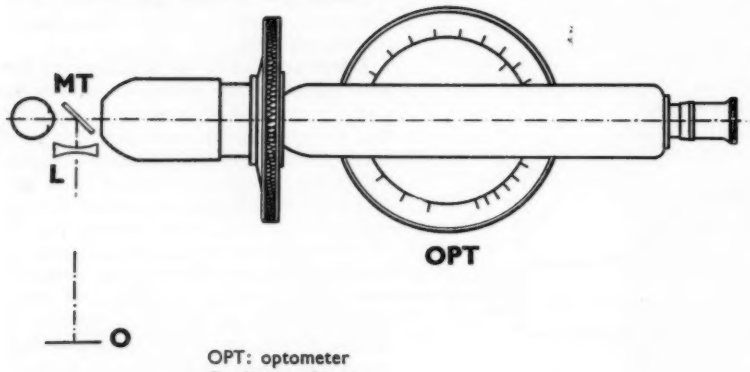
FIG. 2. ACCOMMODATION PRODUCED AS A REFLEX BY ALTERING THE CONVERGENCE OF THE VISUAL AXES



Abscissae: convergence (in reciprocals of the distance to which the eyes converge)
Ordinates: accommodation (in dioptres)

Records are of an exophoric and an esophoric subject.

FIG. 3. DIAGRAM (PLAN VIEW) OF APPARATUS FOR OBSERVING ACCOMMODATION REFLEX TO CHANGES IN THE DIVERGENCE OF LIGHT FROM THE OBJECT



OPT: optometer
O: distant object
L: negative lens (or adjustable telescope)
MT: transparent mirror

3. Chromatic Stimulus to Accommodation

One characteristic of the image which will vary with focusing is caused by the normal chromatic aberration of the eye. The retinal image of a white source of light, or in fact of any object which is not monochromatic, has coloured fringes which change as the focus of the light changes. These fringes surrounding a white object upon a dark field are red when the dioptric power of the eye is too low (hypermetropia) and blue when it is too high (myopia). We may presume that the retina, being colour sensitive, could detect these colour differences and that the brain could interpret them in terms of the direction that the adjustment of accommodation should take.[1]

To test whether the characteristics of the image produced by chromatic aberration provide the stimulus to the reflex, the experiment described above was repeated with monochromatic, instead of white light. In this experiment the object, illuminated by a sodium vapour lamp, was a field of 1° diameter bearing a pattern of black dots about 3′ diameter upon it. It was found that many subjects showed no reaction to changes in the vergence of the light from this monochromatic object, although their reaction with white light was normal. More than 60 subjects have been tested in this way; over 30% of them showed no reaction and 25% gave an abnormal reaction which showed that in these also the true reflex was not operating. The nature of this abnormal reaction is interesting. Owing to the eye being deprived of the information normally given by chromatic aberration, it attempted to make use of the blurring of the image, although the subject made no conscious effort to obtain clear vision. Thus, when the rays of light from the object were made to be divergent at the eye, resulting in a blurred retinal image, the eye accommodated and produced a sharp image, but when the light was again made parallel, the resulting blurred image in monochromatic light gave no information as to which way the mechanism should be adjusted. The result was that, although these subjects would accommodate when necessary, they could not relax accommodation again while viewing the monochromatic object, and they were left with blurred vision until the light was changed to white or other suitable heterochromatic mixture, when the normal reactions were restored.

Evidence that the reaction of positive accommodation in these subjects was stimulated by a blurred image when the light was monochromatic is given by their behaviour when a positive lens was placed before the eye; they distinctly accommodated, although in this case such an adjustment was in the wrong direction, and still further decreased the definition of the image.

4. Achromatic Stimulus to Accommodation

Many subjects, at least 40% of those who have been tested, have been found to react quite normally to changes in the vergence of monochromatic light if the over-all diameter of the object is sufficient. We must infer, therefore, that characteristics of the retinal image, other than those resulting from chromatic aberration, can give the necessary information for the control of accommodation as a reflex. Some experiments which are now in progress show that in these cases, when the object is illuminated with white light, both the chromatic and the achromatic factors are made use of by the mechanism, and their effects appear to be additive. This is demonstrated by varying the size of the fixation object. If a small circular bright object of variable size is used, a minimum diameter for which the reflex still operates can be found. This value is never less than 2′ under the best conditions and more frequently is between 4′ and 10′. If the size of the object is reduced below this minimum, the accommodation does not react to changes in the vergence of the light. Usually the ability to relax from the accommodated condition fails first and the whole reaction becomes slower as the end point is reached. If monochromatic light is used, the minimum size of object which allows the reaction is larger by a few minutes than the minimum required for a white object of equal luminance.

The nature of the achromatic stimulus cannot be stated with certainty at present, but some phenomena associated with it have been recorded which may lead to a solution of the problem. In using the optometer to watch the changes in the dioptric power of the eye, one observes the retinal image of a target located in the instrument. It can be seen that the subject's fixation is not constant: details of the fundus of the eye, such as small retinal vessels, are seen to make constant small oscillating movements with regard to the image. It is estimated that these movements are rather less than 10′ of angle. In most cases subjects cannot prevent these movements, particularly when the vergence of the light from the fixation object is altered. However, a few subjects have been found who can maintain good fixation if the change in the vergence of the light is small. In these cases no accommodation occurred when the divergence of the light from the object was changed, so long as fixation was held, but as soon as the subject, under direction, turned the eye to fixate another mark at an angle of 6′ to 10′ from the first, the eye accommodated. This effect of fixation has been observed for both positive accommodation and relaxation. Another effect that has been recorded is that there is no reaction if vision is directed slightly

[1] See Hartridge, p. 31 of this Bulletin.

away from the object. No very precise figure has been obtained for this, but it has been found that, if vision is allowed to scan around the object but not approach it within 10′, no reaction occurs. This is taken to indicate that only the central region of the fovea is concerned in initiating the reflex. These effects regarding fixation have been found with both white and monochromatic objects.

The existence of the achromatic stimulus to the reflex shows that the retinal mechanism must possess some means of translating differences of focus into differences of brightness, since brightness and colour are the only fundamental differentiations the retina can make. The experiment with fixation leads to the conclusion that the minute scanning movements that were observed are concerned with this process of translation. It can be shown by a simple graphical construction that, when the eye is turned so that an out-of-focus image falls away from the axis, the marginal rays on the two sides of the blur circle will strike the retina with differing degrees of obliquity depending upon whether they are divergent or convergent, i.e., whether the light is focusing in front of or behind the retina. If the former, the boundary ray further from the axis will be more normal to the retina; if the latter, the rays on the side of the blur circle nearer to the axis will be more normal. This difference of obliquity could, because of the Stiles-Crawford effect, produce a difference of brightness stimulus.

If this is the means by which the brain gains its information regarding the direction in which accommodation has to be adjusted, the machinery must be highly complex because it involves the detection of minute differences of brightness and the integration of this information with knowledge of the position of the image. Then from the centre which makes the interpretation there must be a link with the accommodation centre and probably with the nuclei of those nerves which control the scanning movement so that the process of detecting and correcting errors in focus can be continuous.

5. Discussion

The fact that there is a chromatic stimulus to accommodation control has an important bearing on some problems of colour vision. The reactions of accommodation to a chromatic stimulus, which have been studied by objective methods, must be initiated by a response by the visual mechanism to changes in the minute colour fringes of the retinal image caused by chromatic aberration. Since these reactions occur without the agency of the cortical centres of perception, we must presume that the impulses conveying the information are transmitted from the retina to the accommodation centre via the superior colliculus, and that at some stage in this path the signals are translated into a form which gives the correct impulse to the motor centre.

If, in the case of colour defectives, the site of the deficiency were in the retina, we should expect the chromatic stimulus to accommodation to be affected. The employment of an object illuminated with a mixture of colours which a dichromatic subject cannot differentiate would prevent the chromatic clues to change in focus being effective. This would result in the chromatic stimulus being absent under these circumstances. If the subject were in the group which reacts to the achromatic stimulus we should expect to find no difference in the minimum size of object to which the reaction occurs, whether it be illuminated with monochromatic light or with the heterochromatic mixture. If on the other hand a colour-defective subject were in the group which does not show the reaction with monochromatic light, as we have found about 50% of normal trichromats to be, there would be no accommodation reflex at all to changes in the vergence of the light. Such results would lead to the conclusion that a defect in the colour-receiving mechanism existed in the retina at a stage which affected not only the visual impulses to the geniculate body but also those leading to the colliculus and concerned with the initiation of reflexes.

If it were found that the results of the experiments on colour-blind subjects showed that their accommodation reflexes to changes in vergence of light were the same as those in trichromats, we should infer that the retinal impulses which initiate the reflexes were normal and that the defect in the colour mechanism must lie in the geniculate body or in the cortex. As it has been shown that only a relatively small central area of the fovea is concerned in the stimulation of this reaction, these deductions regarding colour anomalies must be limited also to this region.

It may be interesting to speculate on the purpose of this direct reaction of accommodation to stimulation by changes in light vergence, which is independent of the subject's consciousness of nearness of the object. The reflex operates over a relatively small amplitude, probably not more than 2.5 D., but within this it constitutes an automatic focusing device, governed by the actual vergence of the light at the retina. It probably acts as a fine adjustment for correcting inequalities between the functions of accommodation and convergence, for which purpose an amplitude of 2.5 D. would be sufficient. Such inequalities exist in cases of errors of refraction and in the common conditions of imbalance of the ocular muscles. For example, the hypermetrope must accommodate to see distant objects clearly but must maintain parallelism of the visual axes; the myope must exert convergence in excess of accommodation when viewing near objects binocularly. An exophoric subject will need a greater innervation for a given degree of convergence than one who is orthophoric, and, as we have seen, this greater innervation would also be applied to the accommodation mechanism and would cause it to operate in excess of the physical requirement. These mutual inconsistencies are made compatible by the direct reaction of the mechanism to the divergence or convergence of the light rays at the retina.

Another example of the need for control of accommodation independently of convergence arises from the variation in the response of the crystalline lens to forces applied to it by the lens capsule when the ciliary muscle contracts. In the record for the exophoric subject (fig. 2), it is seen that, as the near point is approached, there is a gradual reduction in the progress of accommodation induced by increasing convergence. This gradual reduction of progress is found in all cases and indicates that towards the upper limit of the amplitude of accommodation the innervations which normally accompany convergence are insufficient. Consequently, in order to produce a sharp retinal image of objects at these short distances, relatively greater innervations and greater contractions of the ciliary muscle are required. The control for these innervations must also depend upon the light vergence stimulus.

References

Adamson, J. & Fincham, E. F. (1939) *Trans. ophthal. Soc. U.K.* **59**, 163

Fincham, E. F. (1937) *Proc. phys. Soc. Lond.* **49**, 456

Fincham, E. F. (1951) *Brit. J. Ophthal.* **35**, 381

THE NATURE OF REFRACTIVE ERRORS

ARNOLD SORSBY M.D. F.R.C.S.

Royal College of Surgeons of England

Royal Eye Hospital, London

1 Variations in the optical components
  a Variations in the cornea-lens system
  b Variations in axial length
2 The nature of refractive errors
3 The emmetropic eye
References

Anatomical findings—mainly post mortem—appear to have established that the myopic eye is a long eye, and the hypermetropic eye a short eye. This finding is all the more readily accepted as it is consistent with the optical concept of the emmetropic eye as one in which ingoing parallel rays of light are focused on the retina. In this simple scheme emmetropia, hypermetropia, and myopia are therefore interpreted almost exclusively in terms of axial length, it being held that the emmetropic eye is 24 mm. long, that the hypermetropic eye is shorter than this, and the myopic eye longer. This scheme has been responsible for an immense amount of work which aimed at establishing the factors that lead to myopia as a result of axial elongation.

A number of well-established facts make this simple conception largely untenable. The axial length is only one of several factors that determine the refractive state of the eye. Refraction of rays of light occurs in the anterior segment of the eye—at the cornea-lens system—and it is obvious that an eye may still be emmetropic even if it has a long axis, provided the refraction at the cornea-lens system is low. The simple view of ocular refraction generally held is based on the assumption that the cornea-lens refractive system is a constant. Actually this is not so. Each of the components of refraction shows considerable variation. No understanding of refraction is possible without an appreciation of the correlation of these variable components.

1. Variations in the Optical Components

a. Variations in the Cornea-Lens System

Cornea

The variations in the radius of curvature, and consequently in the refractive power of the cornea, have been extensively studied by many investigators since Steiger in 1908–13 first emphasized this aspect (see Steiger, 1913). On the basis of measurements of the refractive power of the cornea of 5,000 eyes in children, Steiger showed that the measured values ranged from 39.0 D. to 47.0 D., and emphasized that the frequency of distribution followed a theoretical binomial curve. Subsequent investigators have largely confirmed this.

Anterior Chamber

Here, too, there is considerable variation: Tron (1929, 1940), in one series of measurements of the depth of the anterior chamber, found a variation of 2.16 mm. to 5.05 mm. Most values fall between 3.0 mm. and 4.5 mm. Whilst the depth of the anterior chamber does not play a large part in the total refraction of the eye, the considerable variation observed is significant.

Lens

Ever since 1890 when Tscherning introduced his ophthalmophakometer, it has been possible to measure fairly accurately the radii of curvature of the lens and compute its refractive power. Technical difficulties appear to have limited the number of investigations on this important optical component, but it is now established (Awerbach, 1900; Zeeman, 1911; Tron, 1934) that the refractive power of the lens varies from 14 D. to 25 D., and that once more the distribution is binomial.

b. Variations in Axial Length

Computed Values

No direct measurements of axial length were possible until the *x*-ray method of measurement was introduced. Earlier workers could fall back on two different methods:

i. *Measurement of enucleated eyes.* In a series of measurements of 122 enucleated eyes, Erggelet (1936) reported a variation in axial length between 21.7 mm. and 28.75 mm., with an average of 24.68 mm. The refraction of such eyes was not known, and the distribution was held to conform fairly closely to a binomial curve.

ii. *Computation of axial length on the basis of clinical refraction and actual measurement of the cornea-lens system.* Such measurements have been recorded by Awerbach (1900), Zeeman (1911) and Tron (1929, 1930, 1934). The measurements all show considerable departure from binomial distribution if all refractive states are taken into account. When highly myopic eyes are excluded there is a closer approach to binomial distribution, but with a considerable excess for emmetropia.

Axial Length as Determined by x-Ray Measurement

Two series of *x*-ray measurements of axial length are available: a study on a thousand eyes by Stenström (1946), and a small series (of 45 eyes) by Deller, O'Connor & Sorsby (1947). Both series show a considerable variation in the axial length of emmetropes. In Stenström's series, eyes with a refraction of 0 D. to $\pm$ 1.0 D. show axial lengths between 21.5 mm. and 26.5 mm. In the second series, the axial length of the emmetropic eyes ranges between 21.0 mm. and 25.0 mm. Both series likewise show that hypermetropic eyes tend to have a shorter axis, and myopic eyes a longer axis; and both also bear out the fact that most myopic and hypermetropic eyes have an axis that falls within the emmetropic range. It is only in the higher degrees of hypermetropia and of myopia that the axis tends to fall outside this range.

2. The Nature of Refractive Errors

With the three components of the cornea-lens system each showing a wide fluctuation in emmetropia, and the axial length also showing a wide range, the varieties of refractive errors can readily be visualized as the result of the free

combination of the components of total refraction of the eye with variable axial length. Tron (1934) has emphasized this concept in designating most refractive errors as "combination ametropia". This reading is, however, once more too simple, because of the following considerations:

i. Even if the different components of refraction show a binomial distribution, the refractive state of an unselected population shows no such distribution. The many refractive curves that have been established for unselected populations do indeed show some symmetry as between the hypermetropic and the myopic limbs, but the outstanding facts are that there is a marked excess of refractions in the near neighbourhood of emmetropia, and a marked excess of high myopes.

ii. The marked excess of emmetropic and near-emmetropic eyes invalidates any suggestion that there is a free combination of these variables. The eye is obviously a co-ordinated organ and not a haphazard combination of variable components. The concept of combination ametropia may explain why all eyes are not strictly emmetropic, but it can be only an auxiliary process against the major process of correlation of the variable components.

iii. During the process of growth of the eye there is an increase of some 8 mm. in the axial length of the eye, for the eye of the new-born baby is some 16 mm. long, whereas the eye of the adult is somewhere around 24 mm. This axial elongation presupposes that either the new-born baby shows some 30 D. of hypermetropia or, alternatively, if it is nearly emmetropic, that the adult eye shows some 30 D. of myopia. Both these suppositions are absurd, and it is obvious that considerable changes in the refracting surfaces of the cornea and lens must occur during growth to maintain something like an emmetropic refraction. In fact, the growth of the eye is itself a highly correlated process, involving marked changes in the refracting power of the anterior segment of the eye. This once more emphasizes that the eye is a correlated organ and not merely a combination of freely variable components.

iv. Any purely optical explanation of the refractive errors overlooks the well-established clinical experience that most myopia develops and progresses during adolescence, when the eye of the emmetrope is already a finished and stable organ.

These, and such-like considerations, whilst they do not minimize the valuable work on variation of optical components, do emphasize that the primary question is the nature of the correlating mechanism that produces emmetropia. An understanding of the refractive errors can only follow on an understanding of the normal.

3. The Emmetropic Eye

Work which we are now engaged on shows the adult emmetropic eye as possessing the following features:

i. *A refracting cornea-lens system of no constant value.* The range of corneal refraction is from 39.9 D. to 47.2 D.; the depth of the anterior chamber is from 2.7 mm. to 4.2 mm.; the radius of curvature of the anterior surface of the lens is from 9.2 mm. to 12.2 mm., and that of the posterior surface of the lens is from 5.4 mm. to 7.1 mm.

ii. *The axial length.* This likewise has no constant value but extends from 22.2 mm. to 25.8 mm. As the total axial length includes the refracting anterior segment it may be best to consider the reduced axial length from the retina to the posterior surface of the lens; this ranges from 14.9 mm. to 18.3 mm.

iii. *Spherical shape.* In 18 emmetropic eyes the transverse and vertical diameters measured by the *x*-ray method showed that only exceptionally was there any departure from spherical shape. For the series as a whole, the ratios of the transverse diameter and of the vertical diameter to the axial diameter were 1.0 and 0.99, respectively. The computed deviation from spherical shape was 1.00, as determined by the formula $\frac{\text{transverse diameter} \times \text{vertical diameter}}{\text{axial length}^2}$. (The findings for 11 hypermetropic eyes were not markedly dissimilar, but those for 15 myopic eyes showed an actual and relative elongation of the axis, with a mean deviation from spherical shape of 0.90.)

iv. *Correlation of axial length with lens power and corneal power.* In a series of 121 emmetropic eyes the axial length was plotted against the calculated power of the lens, derived from measurement of the refracting surfaces, and against corneal power as determined by keratometry. There was good correlation for each function separately.

These findings all emphasize that the emmetropic eye is a finely correlated organ, and that the actual mechanism of correlation still requires to be elucidated. It must be postulated that this correlating mechanism is effective during childhood when the eye grows, and the problem of the refractive errors becomes reduced to disclosing in the first instance whether there is any failure in correlation of the different components and diameters. Such studies will merely reveal the optical basis of refraction; they cannot by themselves explain the mechanism of faulty correlation. It is likely that faulty correlation is genetically determined; a study of these components in different families is therefore essential. We are carrying out such studies, and also investigations on the changes during the growth of the eyes.

References

Awerbach, M. I. (1900) *The dioptric of eyes with various refraction.* [Thesis, Moscow.] Cited by Tron (1940)

Deller, J. F. P., O'Connor, A. D. & Sorsby, A. (1947) *Proc. roy. Soc.* B. **134,** 456

Erggelet, H. (1936) *Ber. dtsch. ophth. Ges.* **51,** 214

Steiger, A. (1913) *Die Entstehung der sphärischen Refraktionen des menschlichen Auges.* (*Übersichtigkeit—Normalsichtigkeit—Kurzsichtigkeit.*) Karger, Berlin

Stenström, S. (1946) *Acta ophthal., Kbh.* Suppl. No. 26

Tron, E. J. (1929) *v. Graefes Arch. Ophthal.* **122,** 1

Tron, E. J. (1930) *v. Graefes Arch. Ophthal.* **124,** 544

Tron, E. J. (1934) *v. Graefes Arch. Ophthal.* **132,** 182

Tron, E. J. (1940) In: Ridley, F. & Sorsby, A., ed. *Modern trends in ophthalmology*, vol. 1, p. 245. Butterworth, London

Zeeman, W. P. C. (1911) *v. Graefes Arch. Ophthal.* **78,** 93

THE INTERPRETATION OF SPECTRAL SENSITIVITY CURVES

H. J. A. DARTNALL B.Sc. Ph.D. F.R.I.C.

Medical Research Council
Group for Research in the Physiology of Vision
Institute of Ophthalmology, London

1 A visual pigment nomogram
2 Physical factors in the correlation of sensitivity curves with visual pigments
3 Application of the nomogram in the interpretation of some spectral sensitivity curves
4 Discussion
References

One of the objectives of visual research is the quantitative interpretation of the spectral sensitivity curves. A notable, but solitary, success is the relation which has been established between the absorption spectrum of the light-sensitive pigment, visual purple, and the scotopic sensitivity of certain animals, including man.

There are a number of reasons why other correlations of this nature have not been attempted or have met with indifferent success. The main reason is ignorance of the number and distribution of the visual pigments in the animal kingdom. Apart from the nearly ubiquitous visual purple (or rhodopsin), discovered just over one hundred years ago, the only visual pigments known until recently were visual violet (or porphyropsin), found in certain fishes (Köttgen & Abelsdorff, 1896), and iodopsin, found, together with visual purple, in the chicken (Wald, 1937; Bliss, 1946).

Another reason is that, since extracts of the visual pigments from dark-adapted retinas are always impure, their absorption spectra are only imperfectly known. The presence of stable impurities in the extracts can be tolerated in most work since, under selected and controlled conditions, a visual pigment is adequately characterized by the changes in absorption which take place on bleaching it; in other words by the difference between the absorption spectrum of the visual pigment and that of the product of its bleaching. Such " difference " spectra (as they are called) are obtained by subtracting the measured absorption of the bleached from that of the unbleached extract, a procedure which eliminates the effect of impurities. On the other hand, for comparison with spectral sensitivity curves the absorption spectra of the visual pigments themselves must be known. Only for visual purple, however, have methods of extraction been perfected to the point where solutions of near purity are obtained.

Recently, additional visual pigments have been found. These (so called from the respective wavelengths at which they absorb light most strongly) are visual pigment 467, found in the tench (Dartnall, 1952a), visual pigment 510 in the bleak (Dartnall, 1952b) and visual pigment 519 in the clawed toad, *Xenopus laevis* (Dartnall, unpublished).

From the present view-point the main interest of these findings lies not so much in the pigments themselves as in the revelation of some remarkably constant properties. Thus (when drawn with frequencies as abscissae instead of the more usual wavelengths) the difference spectra of alkaline solutions (pH 8.5) of the new pigments and also those of visual purple, visual violet and iodopsin are similar in shape. It is in alkaline solution that the pH-dependent absorption band of the product of bleaching of any visual pigment is located farthest from the absorption band of the parent pigment. Moreover, the separation is sufficiently large for the " positive " portion of the " alkaline difference spectrum " to be an approximation to the true absorption spectrum of the visual pigment. The similarity of shape of all the alkaline difference spectra in their positive portions is most marked, and strongly suggests that the absorption spectra of the visual pigments themselves are closely similar in shape. Consequently, a knowledge of the absorption spectrum of one would suffice to calculate a close approximation to that of any other of given spectral position. Furthermore, the fact that the known visual pigments come from a variety of species is evidence for believing that others, as yet undiscovered, will prove to have similarly shaped absorption spectra.

These facts go some way towards removing the abovementioned obstacles in the way of interpreting spectral sensitivity curves. Thus, using the absorption spectrum of visual purple as a basis, a nomogram can be devised by means of which the absorption spectra of other visual pigments can be obtained—at least, approximately. In this paper such a " visual pigment nomogram " is described and, after a review of the physical principles underlying the correlation of sensitivity curves and visual pigments, its application is illustrated. The interpretation of a number of spectral sensitivity curves found by Granit using an electrophysiological technique is then attempted with its aid. The wide range of applicability of the nomogram brings into focus some difficulties and discrepancies which are tentatively considered in the discussion.

1. A Visual Pigment Nomogram

To construct the nomogram it is necessary to know the absorption spectrum of one of the visual pigments. By far the greatest amount of spectrophotometric work has been carried out with visual purple and most modern workers have prepared solutions of this pigment substantially free from light-absorbing impurities. These impurities absorb most strongly in the blue and violet regions of the spectrum, and such variations as exist in the results of different workers are confined to these regions.

Hecht, Shlaer & Pirenne (1942) recommend the averaged data of Lythgoe (1937), Chase & Haig (1938) and Wald (1938) for the absorption spectrum of visual purple. In this assessment the measurements by Schneider, Goodeve & Lythgoe (1939) and by Goodeve, Lythgoe & Schneider (1942) of the spectral variation of the photosensitivity of visual purple were not taken into account.

Since the spectral photosensitivity curve (photosensitivity = absorption coefficient × quantum efficiency) agrees in form with that of the absorption spectrum over a wide spectral range, it may be concluded that the quantum efficiency is constant. Consequently the variation of photosensitivity with wavelength is a measure of the true absorption spectrum. The measurement of photosensitivity is not affected by the

presence of stable light-absorbing impurities (Dartnall, Goodeve & Lythgoe, 1936, 1938; Goodeve & Wood, 1938) and is thus of especial value in determinations at the short wavelength extreme of the visible spectrum where optical density measurements are invalidated by the contribution of impurities. The results of Goodeve, Lythgoe and Schneider do, in fact, indicate that the mean values quoted by Hecht, Shlaer and Pirenne for this spectral region are too high. Later measurements with purer solutions by Wald (1949) and Collins & Morton (1950) confirm this conclusion.

Recommended values for the absorption spectrum of visual pigment 502 (visual purple) are given in Table I. These are based partly (460–620 mμ.) on measurements made by the author at 20° C. on a solution prepared from the common frog (*Rana temporaria*) and partly (400–460 mμ.) on measurements made by Goodeve, Lythgoe and Schneider (*Rana esculenta*) and by Collins and Morton (*Rana temporaria*). It has been shown that the same visual purple (visual pigment 502) occurs in both species (Dartnall, 1952b) and that its absorption spectrum is independent of pH (Lythgoe, 1937).

TABLE I. ABSORPTION SPECTRUM OF VISUAL PIGMENT 502 AT 20° C.

Wavelength (mμ.)	Absorption coefficient (percentage of maximum)
400	20.5
410	23.8
420	28.2
430	36.8
440	46.0
450	57.0
460	69.0
470	80.2
480	89.5
490	96.7
500	99.9
502	100.0
510	97.6
520	90.4
530	78.6
540	63.2
550	47.3
560	32.6
570	20.5
580	12.5
590	7.4
600	4.3
610	2.5
620	1.5

A nomogram constructed from the data of Table I is shown in fig. 1. A straight-edge placed across the nomogram intersects the scales at three mutually related points. Thus the example given by the interrupted line in fig. 1 shows that for a visual pigment having its absorption maximum at 533 mμ. (" $\lambda_{max.}$ " scale), the absorption coefficient at 480 mμ. (" wavelength " scale) is 63% of the maximum (" absorption coefficient " scale). Values of the absorption coefficient, expressed as percentages of the maximum, at other wavelengths can be obtained by pivoting the straight-edge about 533 on the $\lambda_{max.}$ scale so that it cuts the wavelength scale at the desired points. By reading off values at a number of wavelengths the absorption spectrum of the pigment may be obtained. The absorption spectra of other visual pigments may be constructed in a like manner by pivoting the straight-edge about the relevant point on the $\lambda_{max.}$ scale.

2. Physical Factors in the Correlation of Sensitivity Curves with Visual Pigments

In the correlation of spectral sensitivities with visual pigments it is necessary to consider the physical relationship between the data compared.

The spectral sensitivity is often specified with respect to the energy incident on the cornea. In such cases the data, to be comparable with the light-absorbing properties of a visual pigment, must be recalculated to allow for the effect of the pre-retinal eye media. These, by reflection and absorption, reduce the intensity of light incident on the visual pigment layer. They are allowed for by dividing the sensitivity measurements at each wavelength by the appropriate transmission factors for the eye media.

Further, since the absorption of light is quantized, the sensitivity measurements should be based not on the number of ergs of energy but on the number of quanta incident per second on the pigment layer. The transformation from determinations in terms of energy is effected by dividing each sensitivity measurement by the wavelength at which it was obtained.

In short, the spectral sensitivity curve should refer to a spectrum of equal quantum intensity at the site of absorption by the visual pigment.

Such an ideal curve is then comparable with the function expressing the spectral variation of the light absorbed by the visual pigment in situ. This is not the same as the absorption spectrum and depends on the concentration of the pigment. To calculate this function from the absorption spectrum a knowledge of the optical density of the pigment in the retina is required, the fraction of light of wavelength λ absorbed by a pigment of optical density D_λ being given by $(1-10^{-D_\lambda})$. In this calculation the absorption spectrum data should apply to the temperature at which the sensitivity measurements were made (for cold-blooded animals) and to 37°–42° C., depending on the species, for warm-blooded animals.

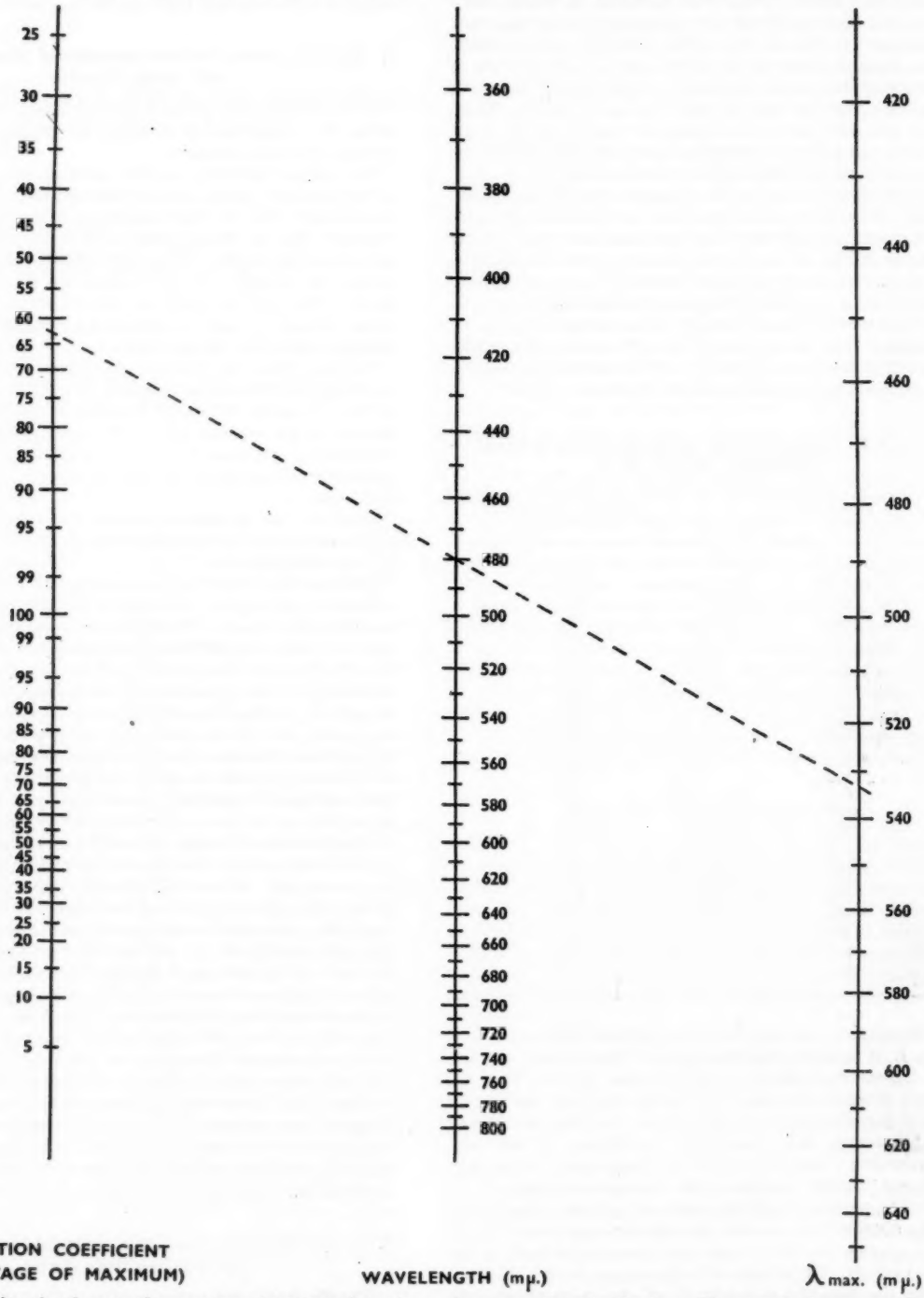
These are the minimum physical considerations. In some cases allowances may also have to be made for reflection from a tapetum (e.g., in the cat), light absorption by coloured oil globules (e.g., in birds) or by other intervening pigments in the retina (e.g., the macular pigment of man). Physical conditions may also be affected by the level of light-adaptation of the eye both by physiological changes such as migration of the epithelial pigment (as occurs in, e.g., the frog) and by purely chemical ones such as alteration of pH or the presence in the light-adapted retina of bleached visual pigments competing with the visual pigments themselves for the light (Dartnall, 1948).

If all these physical factors are known and taken into account, then any remaining discrepancies between the data compared are, presumably, due to physiological factors subsequent to the absorption of light by the visual pigment. A properly conducted comparison may thus throw light on the nature of the physiological mechanisms of vision.

3. Application of the Nomogram in the Interpretation of some Spectral Sensitivity Curves

Usually there are insufficient data for a complete comparison of sensitivity and absorption measurements. For example,

FIG. 1. A NOMOGRAM FOR CALCULATING THE ABSORPTION SPECTRA OF VISUAL PIGMENTS



A straight-edge placed across the nomogram intersects the scales at three mutually related points. The example (interrupted line) shows that visual pigment 533 ("$\lambda_{max.}$" scale) at 480 mμ. ("wavelength" scale) has an absorption coefficient 63% ("absorption coefficient" scale) of that at 533 mμ.

in the well-known correlation between the human scotopic curve and " visual purple," the allowances for absorption by the pre-retinal media are still uncertain; the validity of using the absorption spectrum of frog visual purple (visual pigment 502) for the interpretation of human data is now questionable; and the retinal density of the pigment is unknown.

Thus, although the essential relevance of the comparison is undoubted, it has only a semi-quantitative basis, and " reverse " calculations, such as that of Hecht *et al.* (1942) in which the concentration of scotopic pigment in the human retina was estimated from the discrepancies between the retinal sensitivity curve (measured at 37° C.) and the absorption spectrum of frog visual purple (measured at " room temperature "), are of doubtful value.

From the present view-point, however, the missing information is often unimportant or of predictable effect. Thus (judging by the behaviour of visual purple) the absorption spectra of the visual pigments will not be markedly affected by temperature. Consequently, absorption spectra constructed from the nomogram can be used for comparisons with the spectral sensitivity curves of warm- as well as of cold-blooded animals with little additional error arising from the temperature effect.

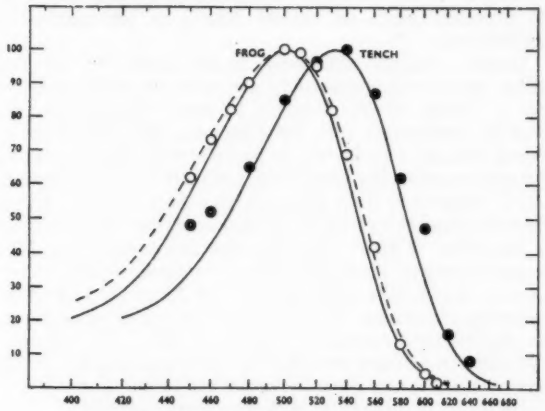
Again, although the curve describing the spectral variation of the light absorbed by a pigment is broader than that of its absorption spectrum, the degree of broadening is only moderate for ordinary concentrations and the position of the maximum is, of course, unchanged. For example, in the case of the frog, shown in fig. 2, the "light absorbed curve" (interrupted line) does not greatly differ from the absorption spectrum (continuous line) even though the former is for an optical density of 0.25 (44% absorption at the maximum)—in the author's experience a high retinal concentration.

In a different category, however, are the corrections for absorption by the pre-retinal media. This absorption is strongest for blue and violet and, because of this spectral variation, has a distorting effect on the sensitivity curve, lowering the apparent sensitivity for short wavelengths more than for long. Moreover, such few measurements as have been made (Sachs, 1891; Roggenbau & Wetthauer, 1927; Ludvigh & McCarthy, 1938) are open to the objection that they were obtained with dead material and also (in the case of human eyes) that this material may have been abnormal. Until measurements with the living eye are perfected these corrections must remain uncertain.

Granit's spectral sensitivity measurements were made with eyes from which the pre-retinal media had been removed in order that the microelectrode could be placed in contact with a point on the inner surface of the retina. A fraction of the vitreous humour probably remained in the optic cup after this operation but since most of the spectral variation in the light absorbed by the media is due to the lens this is of secondary importance. Moreover, reflection losses from the moist retinal surface are nearly uniform throughout the spectrum and may thus be ignored in these relative, as distinct from absolute, measurements. Granit's spectral sensitivity curves, presented for an equal quantum intensity spectrum, are thus already suitable for comparison with the light-absorbing properties of the pigments mediating them. For this reason and because the curves form a survey of a very wide field made with a single technique, the remainder of this paper is devoted to an attempt to interpret them.

Granit (1947) found that most animals investigated yielded

FIG. 2. COMPARISON OF SCOTOPIC DOMINATOR DATA WITH NOMOGRAM CURVES FOR VISUAL PIGMENTS 502 AND 533



Abscissae: wavelength (mμ.)
Ordinates: percentage of maximum

FROG

O Granit's scotopic sensitivity data (private communication)

——— the absorption spectrum of visual pigment 502 (on which the nomogram of fig. 1 is based)

- - - - - - - the spectral variation of light absorbed by visual pigment 502 at an optical density of 0.25

TENCH

● Granit's scotopic sensitivity data (Granit, 1941c)

——— the absorption spectrum of visual pigment 533 calculated from the nomogram of fig. 1

several sensitivity curves according to the state of adaptation of the retina and the position of the exploring microelectrode. Only one type of response, the scotopic dominator, could be elicited from the dark-adapted retina, but the light-adapted retina produced two types: the photopic dominator and several modulators, operative in different spectral regions.

Granit considers that the function of a dominator, a sensory mechanism with a broad spectral sensitivity curve, is to make " a large range of wave-lengths available for vision "; that of the modulators, which have narrow sensitivity curves, the mediation of wavelength discrimination. Thus both the dark- and light-adapted retina have their separate " luminosity " mechanisms (the dominators) while the light-adapted retina possesses, in addition, mechanisms (the modulators) which may act as the peripheral mediators of colour vision provided that a central colour-appreciating faculty exists.

In his examination of dark-adapted retinas, Granit found only two markedly different dominators: one, having a maximum at about 500 mμ., present in the retinas of the cat, rat, guinea-pig, frog, toad, pigeon and eel, and the other a maximum at about 530 mμ., present in the retinas of the tench and carp. He averaged the results from the four species first mentioned and compared the mean spectral sensitivity curve with the absorption spectrum of visual purple (Granit, 1947). The agreement achieved must be regarded as largely fortuitous as the species whose spectral sensitivities were averaged may not have identical " visual purples " (cf. the observations of Collins, Love and Morton (Collins & Morton, 1950; Collins, Love & Morton, 1952) that rat and " cattle " visual

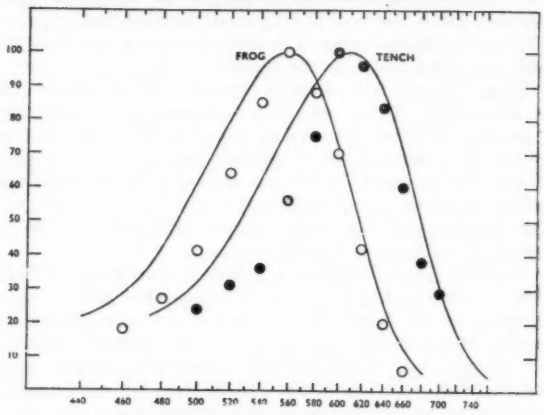
purples have an absorption maximum at 498 mμ., that for the frog being 502 mμ.). Moreover, exact agreement of a sensitivity curve with an absorption spectrum is not to be expected unless the retinal density of the pigment is infinitesimal.

Granit's scotopic dominator measurements for the frog alone (private communication) are shown by open circles in fig. 2. These results, obtained at room temperature, are directly comparable with the light-absorbing properties of visual pigment 502, known to be present in *Rana esculenta*, the species used. The absorption spectrum of this pigment at 20° C. is shown by the continuous line. As would be expected from the probable need for a density correction, the dominator is somewhat broader than the absorption spectrum. The retinal density of visual pigment 502 in *Rana esculenta* is not known, but in *Rana temporaria* it is 0.25 at the absorption maximum (assuming the pigment to be evenly spread over the whole retinal surface) (Dartnall, unpublished). The spectral variation of light absorbed by visual pigment 502 of this density is shown by the interrupted line in fig. 2, which is in distinctly better agreement with the dominator values.

The spectral sensitivity of the other type of scotopic dominator (maximal at about 530 mμ.) is exemplified by the measurements for the tench (fig. 2, filled circles). Although it has long been known (Köttgen & Abelsdorff, 1896) that the tench retina contains visual violet, pure solutions of this pigment have not yet been obtained. Consequently Granit (1941c) could compare his results only with a " difference " spectrum having a maximum at 537 mμ. (Bayliss, Lythgoe & Tansley, 1936). The agreement was only approximate, as was to be expected, but was enough to indicate that the scotopic sensitivity of the tench was mediated by this pigment. The absorption maximum of tench visual violet is now known to be at 533 $\pm$ 2 mμ. (Dartnall, 1952a). Its absorption spectrum may therefore be obtained from the nomogram (fig. 1) and is shown by the continuous line (fig. 2). To a first approximation the agreement between this and the sensitivity data is good and, indeed, since the constructed spectrum may be only a near approximation to the true one, no closer agreement would be expected. It is nevertheless noteworthy that, as in the previous case, the dominator is slightly broader than the constructed spectrum, in accordance perhaps with the need for a " density correction " (for which unfortunately there are no data).

In his microelectrode investigations of retinas in the light-adapted condition Granit likewise found only two distinct dominator mechanisms. The spectral sensitivity curves of these were situated some 60–70 mμ. higher in the spectrum than were those of the corresponding scotopic dominators (Granit, 1941c). Thus the photopic dominators of the cat (Granit, 1943b) and frog (Granit, 1942a) had sensitivity maxima at about 560 mμ. and those of the tench and carp (Granit, 1941c) at about 610 mμ. An exception was the pigeon (Granit, 1942c) which, though having a scotopic dominator with maximum sensitivity at 500 mμ., had a photopic dominator with maximum sensitivity at 580 mμ. However, the pigeon retina contains intensely coloured oil globules, and Granit considered that it was the selective light absorption by these which had displaced the apparent maximum sensitivity of the photopic dominator from 560 mμ. to 580 mμ. In two nearly rod-free species, from which no scotopic response could be elicited, the sensitivity of the photopic dominator was similarly maximal either at 560 mμ., in the grass snake (Granit, 1943a), or at 610 mμ., in the tortoise (Granit, 1941a, 1941c). Conversely, in the rat (Granit, 1941b) and the guinea-pig (Granit, 1942b), two species with few cones, light-adaptation did not produce a Purkinje shift, that is, it did not produce a photopic dominator.

FIG. 3. COMPARISON OF PHOTOPIC DOMINATOR DATA WITH NOMOGRAM CURVES FOR HYPOTHETICAL VISUAL PIGMENTS 560 AND 610



Abscissae: wavelength (mμ.)
Ordinates: percentage of maximum

FROG

O Granit's photopic sensitivity data (Granit, 1942a)

——— the absorption spectrum (calculated from the nomogram of fig. 1) of a hypothetical visual pigment with $\lambda_{max.}$ = 560 mμ.

TENCH

● Granit's photopic sensitivity data (Granit, 1941c)

——— the absorption spectrum (calculated from the nomogram of fig. 1) of a hypothetical visual pigment with $\lambda_{max.}$ = 610 mμ.

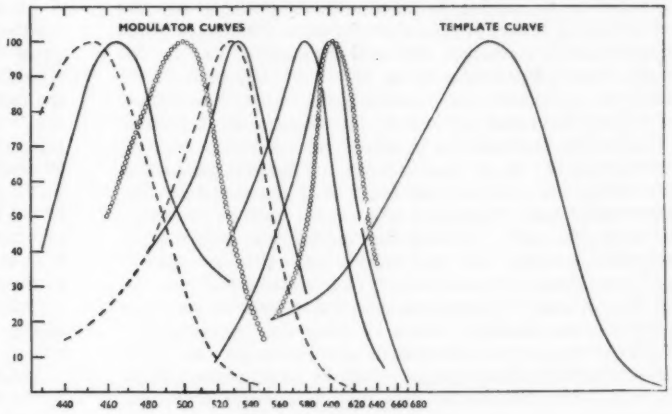
The two photopic dominator types, maximally sensitive at 560 mμ. and 610 mμ., are exemplified by the results for the frog (Granit, 1941c) and the tench (Granit, 1941c) respectively and are shown in fig. 3. Neither in the frog nor in the tench have visual pigments been extracted which could mediate these responses. The curves in fig. 3 are the absorption spectra, calculated from the nomogram of fig. 1, of hypothetical visual pigments having maxima at 560 mμ. and 610 mμ. Both these curves are too broad to fit the sensitivity data. The near fit of the right-hand limbs of the curves is not significant since the maxima of the dominators are not precisely located in the spectrum, and by construction of hypothetical curves with maxima at slightly longer wavelengths the sensitivity data could be more symmetrically straddled.

Thus the spectral sensitivity curves of the photopic dominators cannot be interpreted in terms of single visual pigments. This, however, is not surprising, for Granit (1947) has adduced evidence for the belief that the dominator mechanisms are complex and probably consist of a synthesis of modulator mechanisms.

In addition to the photopic dominator mechanisms, light-adapted retinas also possess modulator mechanisms operative over comparatively narrow spectral ranges. The spectral sensitivity curves of a number of these found in the guinea-pig, rat and frog are shown in fig. 4 (replotted from Granit,

1943c). In this, as in figs. 2 and 3, the horizontal axis, though calibrated in wavelengths (mμ.), is in fact a frequency scale. With such a scale the shapes of the absorption spectra of the visual pigments are identical—or nearly so—with that of the template curve in the extreme right of the figure. The template curve is in tolerable agreement with that of the 450 mμ. guinea-pig modulator but is broader than all the others and much broader than some, particularly those situated in long wavelength regions of the spectrum. Thus the modulator curves cannot be directly correlated with visual pigments—at least not with visual pigments having absorption characteristics similar to those known.

FIG. 4. COMPARISON OF THE MODULATOR CURVES WITH THE TEMPLATE CURVE



Abscissae: wavelength (mμ.)
Ordinates: percentage of maximum

Modulator curves: spectral sensitivity curves found by Granit in light-adapted retinas of the following animals (Granit, 1943c):

- - - - - - - - guinea-pig (two curves maximal at 450 mμ. and 530 mμ.)

o o o o o rat (two curves maximal at 500 mμ. and 600 mμ.)

———— frog (four curves maximal at 465 mμ., 535 mμ., 580 mμ. and 600 mμ.)

Template curve: the common shape of the absorption spectra of all known visual pigments

Note: In this figure (as in figs. 2 and 3) the axis of abscissae, though calibrated in wavelengths for convenience, is a scale of frequencies. This accounts for the progressive diminution in size of the wavelength intervals as the red end of the spectrum is approached. When a scale of frequencies is used for the abscissae the absorption spectra of all visual pigments have the same shape (that of the template curve).

4. Discussion

Does the failure of the template curve to describe the shape of the modulator curves indicate that the latter are due to visual pigments of a different type, characterized by very narrow absorption spectra? Ball *et al.* (1948) have shown that, when vitamin A_1 or retinene$_1$ are dissolved in concentrated sulphuric acid or syrupy phosphoric acid at temperatures near 0° C., coloured solutions are obtained, the absorption spectra of which exhibit a number of well-defined maxima. The difference in rates at which the maxima faded when the solutions were warmed or exposed to light suggested that they contained a number of pigments each possessing a single absorption maximum. The absorption spectra of these pigments were even narrower than Granit's modulator curves. These authors have thus shown that pigments with absorption spectra as narrow as the modulator curves can exist and, moreover, can be prepared from precursors of visual purple, albeit by unphysiological methods.

Against the implication from these experiments that the modulator curves are directly referable to visual pigments is the fact that all known visual pigments have absorption spectra following a broader pattern even though half of them (iodopsin, visual pigment 467 and visual pigment 510) are, presumably, cone pigments. Again, if there were a series of cone pigments with narrow spectra, it might be expected that these, like the known pigments, would have a single shape in common when plotted with frequency abscissae. A glance at fig. 4, however, shows that the breadth of the modulator curves becomes progressively less as their location advances to the red end of the spectrum.

On the other hand, if the sensitivities of the modulators are mediated by " normal " visual pigments having maxima coincident with the peaks of modulator sensitivity, how are the shape discrepancies to be explained ? A possible clue is given by the 500 mμ. rat modulator (fig. 4). Granit (1947) was uncertain whether or not to include this with the other modulators as it differed from them in two important respects. Firstly, it was found only in retinas containing rods and secondly it was less resistant to light, i.e., adaptation affected the shape of its spectral sensitivity curve (though not the position of the maximum). Thus a single active unit, which behaved as the 500 mμ. scotopic dominator when the rat retina was dark-adapted, became modulator-like on light-adapting the retina (Granit, 1941b). The inference is that throughout the experiment the response was due to a rod mechanism utilizing visual purple but that light-adaptation altered its character.

According to the duplicity theory, the rods mediate scotopic and the cones photopic vision. Although in the human, and probably in any, eye the number of fibres in the optic nerve (500,000) is not nearly enough to provide a separate pathway from each of these retinal end-organs (6,600,000 cones, 110,000,000 rods) (Lythgoe, 1938), it is generally considered that the majority serves the cones. Thus there are many more rods per fibre than cones per fibre. In order to account for the variation of visual acuity with the state of retinal adaptation Lythgoe (1940) was led to the postulate that the number of end-organs served by an optic nerve fibre increased (by a spread of synaptic connexions) on dark-adaptation and vice versa.

This amounts to saying that on light-adaptation the relationship of rod to fibre tends to become more like that of the cones. Should this be true, then the fact that light-adaptation causes the response of the rat scotopic dominator to become modulator-like implies that this narrowing of the sensitivity curve is associated with synaptic changes. The narrowness of the other modulator curves (which are presumably mediated by the cones) is then ascribable to the effects of the normal relationship of cone to fibre.

There is, however, a fundamental difficulty in the implica-

tion that mechanisms in the neural layers of the retina can affect the shape of the sensitivity curve of a sensory mechanism. It is difficult to imagine the visual response arising in any way other than as some function of the light absorbed by the visual pigment in the end-organ (Dartnall & Goodeve, 1937). The light absorbed is given by $I_\lambda\ (1-10^{-D\lambda})$, where I_λ is the intensity of incident light of wavelength λ and D_λ is the optical density of the visual pigment for this wavelength. For small values of D_λ, such as occur in retinas, this expression approximates to $I_\lambda D_\lambda$. In the determination of spectral sensitivities the values of I_λ are recorded which elicit a response of constant magnitude. Constancy of response requires constancy of $I_\lambda D_\lambda$, the light absorbed, no matter how complex the relations between them may be. For any invariant value of $I_\lambda D_\lambda$ throughout the spectrum, I_λ will be an inverse measure of D_λ and hence $1/I_\lambda$ will reproduce the absorption spectrum of the visual pigment. Cases in point are the sensitivity curves of the scotopic dominators which reproduce the light-absorbing properties of visual purple or visual violet. These facts make it difficult to understand how synaptic changes could affect the form of the sensitivity curve for the rat scotopic dominator, assuming that only rod mechanisms utilizing visual purple were involved. It is likewise difficult to envisage how the modulator curves could be based on single visual pigments of " normal " properties.

There is, however, a great deal of interconnexion of retinal end-organs of all types through the bipolar, amacrine and ganglion cells (Polyak, 1949). The spectral sensitivity of a combination of different types of end-organ will depend on the relative numbers involved and the nature of the interaction in the neural layers.

As a hypothetical example, a unit consisting of equal numbers of two different end-organs containing respectively equal amounts of visual purple and visual violet may be considered. Further, let it be postulated that the nature of the interaction is such that only the excess of impulses from the " visual violet " end-organs over those from the " visual purple " end-organs is allowed through the neural layers, any inverse effect being inhibited. Then the net spectral sensitivity of the unit would approximate in form to the positive part of the curve obtained by subtracting the absorption spectrum of visual purple from that of visual violet. This curve is of a modulator-like narrowness and is maximal at 565 mμ. Conversely facilitation of the excess of impulses from the " visual purple " over those of the " visual violet " end-organs would yield a modulator-like sensitivity curve maximal at 470 mμ.

Arrangements such as these could be flexible. Variation of the relative numbers of only two types of end-organ (for example, by the existence of different " collecting and editing " stations in the neural layers) would produce mechanisms sensitive in a number of spectral regions. Such flexibility could account for the observed variability of modulator response.

Whether such systems exist is, of course, unknown. At present it would be mere speculation to attempt to account quantitatively for the modulator curves in this way, at least until the visual pigment complements of the retinas concerned are known.

References

Ball, S., Collins, F. D., Morton, R. A. & Stubbs, A. L. (1948) *Nature, Lond.* **161,** 424

Bayliss, L. E., Lythgoe, R. J. & Tansley, K. (1936) *Proc. roy. Soc.* B, **120,** 95

Bliss, A. F. (1946) *J. gen. Physiol.* **29,** 277

Chase, A. M. & Haig, C. (1938) *J. gen. Physiol.* **21,** 411

Collins, F. D., Love, R. M. & Morton, R. A. (1952) *Biochem. J.* **51,** 292

Collins, F. D. & Morton, R. A. (1950) *Biochem. J.* **47,** 3

Dartnall, H. J. A. (1948) *Brit. J. Ophthal.* **32,** 793

Dartnall, H. J. A. (1952a) *J. Physiol.* **116,** 257

Dartnall, H. J. A. (1952b) *J. Physiol.* **117,** 57*P*

Dartnall, H. J. A. & Goodeve, C. F. (1937) *Nature, Lond.* **139,** 409

Dartnall, H. J. A., Goodeve, C. F. & Lythgoe, R. J. (1936) *Proc. roy. Soc.* A, **156,** 158

Dartnall, H. J. A., Goodeve, C. F. & Lythgoe, R. J. (1938) *Proc. roy. Soc.* A, **164,** 216

Goodeve, C. F., Lythgoe, R. J. & Schneider, E. E. (1942) *Proc. roy. Soc.* B, **130,** 380

Goodeve, C. F. & Wood, L. J. (1938) *Proc. roy. Soc.* A, **166,** 342

Granit, R. (1941a) *Acta physiol. scand.* **1,** 386

Granit, R. (1941b) *Acta physiol. scand.* **2,** 93

Granit, R. (1941c) *Acta physiol. scand.* **2,** 334

Granit, R. (1942a) *Acta physiol. scand.* **3,** 137

Granit, R. (1942b) *Acta physiol. scand.* **3,** 318

Granit, R. (1942c) *Acta physiol. scand.* **4,** 118

Granit, R. (1943a) *Acta physiol. scand.* **5,** 108

Granit, R. (1943b) *Acta physiol. scand.* **5,** 219

Granit, R. (1943c) *Nature, Lond.* **151,** 11

Granit, R. (1947) *Sensory mechanisms of the retina: with an appendix on electroretinography.* Cumberlege, London

Hecht, S., Shlaer, S. & Pirenne, M. H. (1942) *J. gen. Physiol.* **25,** 819

Köttgen, E. & Abelsdorff, G. (1896) *Z. Psychol. Physiol. Sinnesorg.* **12,** 161

Ludvigh, E. & McCarthy, E. F. (1938) *Arch. ophthal., Chicago,* **20,** 37

Lythgoe, R. J. (1937) *J. Physiol.* **89,** 331

Lythgoe, R. J. (1938) *Proc. phys. Soc.* **50,** 321

Lythgoe, R. J. (1940) *Brit. J. Ophthal.* **24,** 21

Polyak, S. (1949) *Docum. ophthal.* **3,** 24

Roggenbau, C. & Wetthauer, A. (1927) *Klin. Mbl. Augenheilk.* **78,** 762; **79,** 456

Sachs, M. (1891) Cited by Parsons, J. H. (1915) in : *An introduction to the study of colour vision,* p. 40. University Press, Cambridge

Schneider, E. E., Goodeve, C. F. & Lythgoe, R.J. (1939) *Proc. roy. Soc.* A, **170,** 102

Wald, G. (1937) *Nature, Lond.* **140,** 545

Wald, G. (1938) *J. gen. Physiol.* **21,** 795

Wald, G. (1949) *Docum. ophthal.* **3,** 94

CHROMATIC ABERRATION AND THE POLYCHROMATIC THEORY

HAMILTON HARTRIDGE M.A. M.D. Sc.D. F.R.S.

Emeritus Professor of Physiology, University of London

Gresham Professor of Physic, Gresham College, London

1 The chromatic aberration of the eye
2 Foveal tritanopia
3 The effect of light intensity and visual angle on colour appreciation
4 The nature of the correcting mechanism
5 The polychromatic theory
6 Summary
References

Few problems are more elusive than those which are associated with the resolving power of the human eye. Many disadvantages are apparently present and yet results of outstanding excellence are produced. Thus the lens system of the eye is known to be affected by chromatic aberration and in consequence the images of external objects should be bordered by coloured fringes. But these are seldom if ever observed when the eye is being used in everyday life. In fact so colour-free are these images that Euler (1765) concluded that the lens system of the eye is in some way corrected for this aberration. This caused him to persuade the optician Dollond (1753) to investigate ways in which this correction could be carried out. The result was the discovery of achromatization and the manufacture of the achromatic doublet, consisting of a combination of two lenses, a positive one of crown glass and a negative one of flint glass. This discovery, as is well known, led to advances in optical design of very great importance during the last century. Without these, the microscope and telescope would probably still be in their infancy.

1. The Chromatic Aberration of the Eye

Anatomical dissection discloses the important fact that the lens system of the eye is not corrected in the same manner as the lenses used in these instruments. There are convergent optical arrangements in the cornea and the crystalline lens, but no divergent one, of large dispersion, such as is present in the achromatic doublet. Moreover, there are three experimental conditions which readily demonstrate the presence of chromatic aberration in the eye. These are: (i) making the light rays which enter the eye asymmetrical by obscuring part (usually half) of the pupil, (ii) preventing the light rays from forming a sharp focus by using the eye outside its far point or inside its near point, and (iii) excluding the light rays which correspond to the middle part of the spectrum by means of cobalt blue glass or a purple colour filter. Lastly, measurements of the amount of chromatic aberration present in the eye that were performed by Ivanoff (1947) and by Hartridge (1947) are in good agreement with one another. The amount is closely similar to that present in a crown glass lens of the same aperture and focal length as the eye. Since ordinary crown glass lenses can be achromatized by means of a combination containing flint glass, the same processes should be applicable to the eye. To what extent is vision improved when this correction has been made? Von Helmholtz (1909) was the first to put this idea to the test and he found that placing a correcting lens in front of the eye produced no noticeable improvement in visual acuity. Experiments performed recently by the writer have confirmed von Helmholtz's surprising conclusion. Now there is obviously another way in which the eye can be temporarily achromatized, that is, by replacing ordinary white light by monochromatic light comprising rays of a single wavelength only. This was first done by Luckiesh, who used sodium light for purposes of illumination. He found a small improvement in visual acuity, but this was probably within his experimental error. In his book, *The science of seeing*, Luckiesh states that sodium light had a decided superiority over white light (Luckiesh & Moss, 1937). It is not easy to correlate this statement with the fact that the eye when using white light achieves a resolving power which approximates to the theoretical limit, because it would mean that when using sodium light the eye actually exceeds the limit set by the wave theory of light. This would be most unlikely if not impossible. Moreover, similar tests by Shlaer, Smith & Chase (1942) and by Hartridge (1947) have led to the conclusion that visual acuity is not appreciably improved when white light is replaced by monochromatic light. It has long been known that the resolving power of the eye when grating test objects are used is approximately equal to that calculated for a perfect optical instrument on the basis of the diffraction theory of Ernst Abbe, provided the pupil does not exceed 4 mm. diameter. If the pupil is larger than this, the experimental value of the acuity is found to be less than the calculated value, probably owing to the graininess of the retina, which results from the presence of rods and cones. Recently some experiments have been performed by the writer (to be reported in full elsewhere) in which an artificial pupil of adjustable width has been mounted in front of the eye and then the pupil varied in size until a given grating test object could just be resolved. The tests were then repeated, the same adjustable artificial pupil being used but mounted in front of a well-corrected cine-camera lens with a focal length approximately equal to that of the eye. The same grating test object was used, and the pupil of the camera lens was adjusted until the test object could just be resolved. This was tested in two ways: by photography and by direct visual observation. The photographs were taken on process plates of exceptionally fine grain, which were then examined by means of a low-power microscope. The direct visual observation was carried out by magnifying the aerial image produced by the camera lens by means of an achromatic triplet of magnification $\times$ 10 which was suitably placed in front of the eye. It was found that the width of the artificial pupil was substantially the same in the two cases, showing that, for equal pupils, the eye has approximately the same resolving power as that of a well-corrected cine-camera lens. Thus the chromatic aberration present in the lens system of the eye is corrected in such a manner that resolving power is not in any way interfered with.

We may summarize the position with regard to chromatic aberration as follows: the lens system of the eye suffers from chromatic aberration which is not corrected by optical means in the way that is done in optical instruments. But the coloured fringes produced by chromatic aberration are not

seen in everyday life and hence some non-optical correcting mechanism must be present. This functions in such a manner that the resolving power of the eye is not adversely affected and is equal to the theoretical value calculated on the basis of the wave theory of light.

2. Foveal Tritanopia

Tritanopia or blue-blindness is found in two different forms: congenital and acquired. In the former, there is some defect of colour vision. In the latter, yellow pigment stains the ocular media, particularly the crystalline lens. This pigmentation usually occurs late in life. A given case may exhibit both conditions at the same time. In the congenital type there are two neutral points in the spectrum, one in the yellow and one in the blue-violet.

Closely resembling congenital tritanopia is the so-called "tritanopia of the fovea" that is found when coloured test objects are used which subtend a small angle at the eye. In this form of tritanopia the neutral point is at 0.578 μ. in the yellow; and in the blue-violet the neutral point is at 0.41 μ. according to Willmer & Wright (1945), but at a somewhat longer wavelength according to Hartridge (1950). In support of this conclusion, Richardson-Robinson's case of tritanopia cited by Göthlin (1943) may be referred to:

wavelengths 0.775–0.590 μ. seen as red,
wavelengths 0.589–0.577 μ. seen as neutral,
wavelengths 0.576–0.490 μ. seen as green,
wavelengths 0.489–0.478 μ. seen as neutral,
wavelengths 0.477–0.416 μ. seen as red.

In between the two neutral points the spectrum looks green, outside the two neutral points it looks red. In acquired tritanopia there is no true neutral point at all since all parts of the spectrum that are not absorbed by pigment are seen normally. Observers who have normal colour vision can easily observe the effects of this type of blue-blindness by placing suitable yellow colour filters in front of their eyes. If a powerful beam of monochromatic blue light at a wavelength of about 0.42 μ. be used as a test object, the congenital type of tritanope will identify it as red; whereas, if it is bright enough to be seen by the acquired type, it will be identified as blue or violet.

König (1894), as the result of colour mixture experiments, noted the presence of tritanopia in the normal human fovea. He found, when using small test fields, that all parts of the spectrum could be matched by using mixtures of red rays and blue rays alone. Willmer & Wright (1945) have recently confirmed König's observation and Walls & Mathews (1952) have suggested that foveal tritanopia is due to a non-uniform distribution of colour receptors in the retina, there being too few blue receptors at the fovea. That the acquired type of tritanopia may also be present in normal eyes and be caused by macular pigment has been suggested by Pirenne (personal communication), and he has also suggested that this may assist in the reduction of chromatic aberration by the shortening of the blue end of the spectrum.

Let us then consider the effects which tritanopia of both varieties may be expected to have on the colour fringes produced by chromatic aberration. The first point to be considered is the nature of these fringes. By removing the colour-corrected objective of a telescope and by substituting a simple crown glass lens of the same focal length, the fringes may be observed directly. An even better method is to use a microscope, below the stage of which has been placed an unachromatized Abbe substage condenser of short focal length. The fringes produced by this lens, of external objects reflected in the mirror of the microscope, are then examined by looking down the eyepiece of the instrument. The microscope objective and eyepiece should produce a combined magnification of between ten and forty times. With either of these arrangements the fringes should be readily seen and they are noticed to vary in appearance as the focus is altered. Good definition is usually obtained when the yellow rays are sharply in focus. The fringes at the edge separating a black and a white object are then found to be pale yellow-orange on the white side of the edge and dark blue on the black side of the edge. Somewhat better definition is obtained if daylight illumination is in use when the yellow-green rays are in focus, and in this case the fringes are yellow and blue-violet. When the green rays are in focus the fringes are yellow-green and purple. When the orange rays are in focus they are orange-red and greenish blue.

What colours are the fringes which are normally formed on the retina? In order to decide this point it is necessary to know which rays are most sharply in focus on the retina. The writer used the following method. A grating test object was prepared, consisting of fine parallel strands of unspun silk which were numbered from one onwards. Those having an even number were illuminated by means of electric light from a tungsten lamp, whereas those with an odd number were illuminated by light from a particular part of the spectrum obtained by means of a colour filter having a narrow band transmission. When yellow rays were used both odd and even threads were simultaneously in focus; when rays from other parts of the spectrum were used this was not the case. When daylight illuminated the threads of even number, the threads of odd number were found to be simultaneously in focus with the even ones when they were illuminated through a colour filter which transmitted yellow-green rays. The fringes are then orange-yellow and blue by artificial light and yellow and blue-violet by daylight. On performing spectral analysis on the fringes, which were produced by an unachromatized crown glass lens, it was found that the blue fringe consists of red, green, blue-green, blue and violet rays, while the yellow fringe comprises rays from all parts of the spectrum, with a partial deficiency of red, green, blue-green, blue and violet rays. Thus the blue fringe is spectroscopically purer than the yellow fringe. What effect would macular pigment be expected to have on the fringes? It is usually described as having a pale lemon-yellow colour and it would therefore be expected to absorb blue and violet rays to some extent. This absorption will leave unaltered the blue-green, the green and the red rays present in the blue fringe. It will have hardly any effect at all on the yellow fringe. Similar considerations apply to the effect of foveal tritanopia, supposing it to be due to the absence of blue receptors. In this case the blue fringe will be modified as follows—red will be normal and so will green, blue-green will appear green, blue will appear very dark green, violet will appear very dark red, the yellow fringe will undergo slight changes of the same kind. If, on the contrary, the tritanopia be of the variety seen in congenital blue-blind subjects, then the effects will be much more marked—the red of the blue fringe will appear red and so will violet, blue will appear dark neutral grey, and green and blue-green will appear to be blue-green. The effect on the blue fringe as a whole will be largely to remove its colour and to replace

it by neutral grey. The effect on the yellow fringe will be similar, since yellow is replaced by neutral grey in congenital tritanopia and thus the colour of the fringe will be largely lost. Does experiment confirm these colour losses?

3. The Effect of Light Intensity and Visual Angle on Colour Appreciation

Experiments have been performed in the past using both spectral and non-spectral colours: spectral colours as seen in spectroscopes, isolated by spectral illuminators, or separated by colour filters from the radiations of sodium, mercury or cadmium lamps; non-spectral colours as produced by colour filters, coloured glasses, dyes in solution or applied to papers. It is possible here only to refer to certain parts of this large mass of research work. The effects of reducing light intensity on the spectral colours as seen by the fovea may be described as follows: as the intensity falls the spectrum shortens, particularly at the violet end; yellow gets less and less saturated and ultimately becomes grey, orange changes to red and ultimately to crimson, green changes to blue-green and ultimately to greenish blue. Reduction of light intensity also produces marked alterations in the apparent brightness of the different parts of the spectrum. The luminosity curve of a bright spectrum is fairly smooth, rising to a crest at about 0.55 μ.; but as light intensity is reduced, two parts suffer to a greater extent than the rest, namely, the blue and the yellow; the deficiency of the latter is sharply localized (Sloan's notch), the maximum loss being at about 0.578 μ. (Sloan, 1928). It will be noticed that the colours which suffer the greater losses are those which play a major part in the fringes produced by chromatic aberration.

Reduction of visual angle produces similar effects to those produced by lowering the light intensity: orange changes to crimson, yellow loses both colour and brightness, the latter being localized (Wright's notch), the greatest loss occurring at about 0.578 μ. (Wright, 1946); green is replaced by greenish blue and blue suffers a loss of both colour and intensity. Experiments with dyes and coloured papers confirm the above description; thus a yellow paper is confused with a pale grey paper when either the light intensity or the visual angle is reduced. Normally they are readily separated, owing to the recognition of the presence or absence of blue light, since pale grey paper reflects blue whereas yellow paper does not. But when light intensity or visual angle is reduced, blue is no longer recognized as blue and, in consequence, confusion occurs between the two papers. Now it is clear that neither colour filters, nor coloured papers, still less monochromatic lights, represent the coloured fringes which are produced by chromatic aberration. The above experiments were therefore repeated on the fringes produced by a simple crown glass lens of about 15 mm. focal length, which were viewed by means of several achromatic triplet lenses of different magnifying power. With magnifications of ten times or more, the fringes, both yellow and blue, were clearly visible. With magnifications of six times or less they were invisible. With a magnification of eight times they were just visible. Thus these tests prove the point that the visibility of the fringes is a function of visual angle. The effects of varying the light intensity, with a magnification of ten times, were then investigated by the use of neutral tint filters. With the full illumination of about 100 ft.c. the fringes were clearly seen. At 5 ft.c.[1] they were no longer visible. At 25 ft.c. they could just be seen. It is concluded that variations, both of visual angle and of light intensity, affect the visibility of the yellow and blue fringes which are produced by chromatic aberration, just as they do other colours of the same hue. Moreover, these two variations affect these coloured fringes in the same ways; in fact it can be demonstrated that for every visual angle there is a corresponding light intensity at which the visibility of the fringes is arrested. In practice this works in the following way: when light intensity is high, the fringes would be visible if it were not for the fact that the pupil has contracted and has thus reduced the fringes in width; on the other hand, when light intensity is low, the pupil dilates and the fringes correspondingly increase in width, but this does not bring about their visibility, because of the reduced light intensity. Under practically all conditions there is a surplus correcting effect in excess of that required for purposes of fringe elimination. This may be used for the colour correction of unachromatized lenses used in conjunction with the eye, for example, spectacle lenses, hand magnifiers, eyepiece lenses of microscopes and telescopes, and the like.

Two experiments may be quoted. Some unachromatized crown glass hand magnifying lenses of various focal lengths were tested in conjunction with the eye in average daylight. If the lenses were of longer focal length than 20 mm., no colours attributable to chromatic aberration were noticed. If, however, a lens of half this focal length was used, the fringes were at once perceived. In a second test, four crown glass lenses, two of 10 D. and two of 9 D., were mounted on a board so that they formed an erecting lens system of unit magnification. When distant objects were looked at through this arrangement, no coloured fringes were noticed. Now in such a scheme the aberrations of the four lenses are additive, yet the eye was able to correct the total amount produced. When lenses of double the above focal length were used, then the surplus correction of the eye proved to be insufficient and coloured fringes were found to be visible.

4. The Nature of the Correcting Mechanism

Hypothesis 1. It has been suggested that the fringes of chromatic aberration are corrected because a certain part of the retina, for example, the macula, the fovea or the rod-free area, possesses a reduced form of colour vision quite different from that present in the surrounding retina. Thus the " tritanopia of the fovea " is sometimes mentioned. Would such an arrangement fit the observed facts? This would not be the case for two reasons: by varying light intensity and keeping visual angle constant, it can be shown that the centre of the retina has normal colour vision when light intensity is high, and reduced colour vision when it is low; by performing similar tests at other parts of the retina, it can be shown that they all resemble the centre of the retina in this respect. Moreover, what would be the use of an arrangement which eliminated the chromatic fringes at the centre of the retina and left them uncorrected everywhere else? It may be that the surplus correcting effect described in section 3 is larger at the fovea than it is elsewhere, in some persons, and it may be the reverse in others. Further experimental work is required on this point.

Hypothesis 2. It has been suggested that the receptors for blue light, possibly because they are less numerous than receptors of other types, are affected to a marked extent by

[1] 1 foot candle (ft. c.) = 10.8 metre candles.

variations of light intensity. When illumination is good, they function in a normal manner; when it is not, they become inactive. Such an hypothesis would not of course account for the loss both of hue and of luminosity suffered by yellow, but it would account for these losses in the case of blue if it were not for the important fact that light intensity is only one of the factors which decides whether blue shall be seen or not. There is at least one other factor—visual angle; and there may be a third—retinal adaptation.

Hypothesis 3. There is a mechanism that is probably of a nervous nature, which takes into account both visual angle and light intensity, and possibly also adaptation. How would such a mechanism function? We know that the visibility of a white object on a black background depends on the average solid angle which the object subtends at the eye and on the intensity of illumination. Could not the same principle be applied to the visibilities of yellow and blue without at the same time affecting those of other colours? For such selective action it would seem to be essential that there should be at least four independent nervous pathways which connect the retina with the brain: (i) for red, (ii) for yellow, (iii) for green and (iv) for blue. If conclusive evidence were to be found as the result of experiment that just as yellow suffers loss of both hue and luminosity without red or green being similarly affected, so also blue suffers a loss of both hue and luminosity without either green or violet being affected, then there would have to be a fifth independent nervous path (v) for violet.

5. The Polychromatic Theory

The polychromatic theory postulates that there are present in the human retina for purposes of colour vision several different types of receptor. At the fovea, with light of medium intensities and fields of two degrees, three of these types play the major part. They may be called " orange ", " green " and " blue " because these are the parts of the spectrum at which each type shows the greatest activity. At the fovea, with very high light intensities indeed, two types play the major part, the " yellow " and the " blue ". At the fovea, with low light intensities or small visual angles or with both, two types play the major part, the " crimson " and the " blue-green ". The crimson receptors have activity at the two ends of the spectrum—in the red and in the blue. Now these two components of the crimson receptor are both complementary in hue to the blue-green receptor which acts in conjunction with them. In consequence, there are two neutral regions where these overlap with one another. One of these occurs in the yellow, the other in the blue, parts of the spectrum. Thus they provide the type of colour vision found in the congenital tritanope who describes these two neutral points and finds red at the ends of the spectrum and green in the middle, between the two neutral points. Moreover, this is the kind of colour vision found in those parts of the retina where the fringes of chromatic aberration have been eliminated. At the periphery of the retina, two types play the major part, the " yellow " and the " blue ". In general these different types of receptor resemble Granit's modulators in their spectral distribution (Granit, 1947). Each type of receptor has a pigment associated with it, which resembles those found by Dartnall in animals. It should be pointed out when the corresponding curves of Granit's modulators and the visual pigments found in animals are compared that the former are usually somewhat narrower than the latter. Some possible explanations of this discrepancy will be found in Dartnall's paper in this symposium.[2] Each type of receptor has its own separate nervous connexion with the brain. Consider then what happens when, as a result of chromatic aberration, coloured fringes are produced. Because of the small size of the fringes on the retina, the nervous impulses which are passing to the brain from the yellow and the blue receptors (which are being stimulated by the fringes) either do not get through to the colour-perceiving centres or are suppressed in some other way. The result is that yellow and blue lose both hue and luminosity, causing Sloan's and Wright's notches, and the colours of the fringes disappear (Sloan, 1928; Wright, 1946). If the fringes are not yellow and blue, then the mechanism fails to deal with them adequately. Thus if the fringes be red and blue as seen through cobalt blue glass, or reddish orange and bluish green as seen when the pupil is half covered or the eye is out of focus, then it is unlikely that the mechanism will be able to suppress them. This theory is in satisfactory agreement with the Bezold-Brücke phenomenon as investigated by Purdy (1931), with the photochromatic interval of Abney & Festing (1892), and with the effects of light intensity and visual angle on the visibilities of signal lights. It is also in agreement with other phenomena of colour vision, such as hue and saturation discrimination, the retinal direction effects on hue, and specific adaptation. But it has been suggested by Thomson (1952) that this theory is incompatible with the trichromacy of the colour vision of the fovea with a two-degree field. This point will now be investigated. It is well known that, with the exception of the spectral colours, practically all other colours can be matched by suitable mixtures of red, green and blue rays. Thus if, corresponding to red rays there are receptors for red rays (red receptors) and similarly for green and for blue, which send impulses up three different nervous pathways to the brain for purposes of colour perception, then there is a mechanism for colour vision which superficially fits the facts so far mentioned. It does not of course exclude the possibility that there are also other types of colour receptor in addition to those for red, green and blue. Now the spectral colours cannot be matched by mixing three colours even when the latter are themselves isolated from the spectrum, because in every case the spectral colour is more saturated than the mixture. Thus spectral yellow rays are more saturated than a mixture of spectral red rays and spectral green rays, and this fact suggests the idea that there are yellow receptors in addition to red receptors and green receptors. Now the difference in saturation of spectral yellow and the red-green mixture is not large and this suggests that the yellow receptors are in the minority or in some other manner are less energetic than the red and the green receptors. Similar reasoning applies to spectral blue-green, which is found to be more saturated than a mixture of green and blue, and in this case also there appear to be blue-green receptors which are less active than the green and the blue receptors but are none the less important. Thus human colour vision may be described as being " nearly trichromatic but not quite ".

The hue discrimination of subjects who have normal colour vision discloses the same arrangement: with three types of receptor—red, green and blue—playing the major part. But these alone are found to be insufficient, and in this case also accessory receptors of other types are found to be necessary. We do not know the relative activities of the three main types

[2] See Dartnall, p. 24 of this Bulletin.

and the subsidiary types. They cannot be nearly equal in activity because, if they were, the greater part of the hue discrimination curve would be a straight line. Nor can they be very greatly different because, in this case, the hue discrimination curve would show much more deeply marked undulations. For purposes of tentative calculation the assumption will be made here that the accessory types have one-tenth the activity of the major types for average light intensities and a two-degree field. It will also be assumed that the response curves of the human receptors resemble those found by Granit (1947), namely, the modulators of the retinas of animals; also that the crests of the response curves correspond in wavelength approximately to those of animals. Thus there will be very marked overlap between the responses of one type of receptor and those lying on either side of it in the spectrum. There is already a considerable amount of evidence to justify this assumption. Thus positions in the spectrum of receptors of different type have been investigated in several ways and the results are found to be in good agreement with one another. They are also in agreement with the positions in the spectrum of the absorption bands of the pigments found by Dartnall (1950) in the eyes of certain fish. From information at present available it seems likely that the response curves of receptors adjacent in the spectrum intersect at values of between 60% and 70%. Moreover, if one type be stimulated by a monochromatic radiation which corresponds exactly to the crest of its response curve, then the stimulus produced by the same radiation on the type next to it in the spectrum will be about 25%, and the type next again about 5%, and the next again 1%. Thus, if a monochromatic radiation corresponding in wavelength to the crest of the " orange " receptors stimulated them to a value of 100 units, it would stimulate the adjacent type, namely, the " yellow", to a value of about 25 units, that is, if both had the same sensitivity. But such is not the case, and on the assumption made above of one-tenth the activity for the accessory receptors, the value will be not 25 but 2.5. The green receptors, being major types, will be stimulated to the full value, that is, 5%, and the blue-green receptors, being accessory types, to one-tenth of 1%. On this basis the intensity of stimulation given by a mixture of red rays and green rays has been compared with that given by a pure yellow radiation. The values obtained are given in Table I.

TABLE I. RELATIVE AMOUNTS OF STIMULATION PRODUCED BY A RED-GREEN MIXTURE AND BY A PURE YELLOW

	Red receptors	Orange receptors	Yellow receptors	Green receptors	Blue-green receptors
Red stimulus alone . .	36	91	2	3	—
Green stimulus alone . .	—	5	2	100	2
Red-green mixture . .	36	96	4	103	2
Yellow stimulus alone . .	2	100	40	100	2

If the last two lines of Table I be compared, it will be seen that there are large differences in the cases of the red and the yellow receptors and small differences in the cases of the orange and green ones. The differences suggest that the mixture is too red and too unsaturated for a precise match. The excess redness could be avoided by the use of a somewhat less strong red stimulus in the mixture. The lack of saturation in the mixture is unavoidable. It will be recollected that spectral mixture experiments disclose the same lack of saturation in the case of the red-green mixture.

In Table II are shown the results of a similar comparison in the case of blue-green.

TABLE II. RELATIVE AMOUNTS OF STIMULATION PRODUCED BY A BLUE-GREEN MIXTURE AND BY A PURE BLUE-GREEN

	Yellow receptors	Green receptors	Blue-green receptors	Greenish-blue receptors	Blue receptors	Violet receptors
Green stimulus alone . .	3	133	3	1	1	—
Blue stimulus alone . .	—	1	1	3	133	3
Blue-green mixture . .	3	134	4	4	134	3
Blue-green stimulus alone .	2	100	42	42	100	2

When the last two lines are compared, it is seen that the mixture stimulates the green and the blue receptors more strongly, but the blue-green and the greenish-blue receptors less strongly than does the pure colour. Thus the mixture is less saturated than the pure colour and this conclusion is in conformity with spectral mixture experiments. Calculations similar to these should of course be performed when more precise data are available, but nothing has so far been found to confirm the view that the polychromatic theory is incompatible with the results of spectral mixture experiments.

The theory outlined here and the data shown in Tables I and II imply that six conditions must be satisfied in colour-matching experiments. Is this really the case? When, using a light-adapted fovea and a 2° field, it is required to match spectral colours of longer wavelength than 0.65 μ. by means of mixtures of spectral red (0.65 μ.), spectral green (0.53 μ.) and spectral blue (0.46 μ.), it is found to be necessary to desaturate the spectral red field by adding spectral green to it. Now adding green to the spectral red field is physiologically equivalent to adding the complementary colour to green to the three-colour field, which is adjusted to match the spectral red. In order to match spectral orange or spectral yellow it is similarly necessary either to add spectral blue to the spectral yellow or orange so as to desaturate them, or to add the complementary to spectral blue to the three-colour matching field so as to increase its saturation to that of the monochromatic spectral colour. In order to match spectral blue-green it is similarly necessary either to reduce its saturation by adding spectral red to it, or to improve the saturation of the three-colour matching field by adding to it the complementary to spectral red. Lastly, to match the colours of wavelength shorter than 0.46 μ. it is necessary either to desaturate them by adding spectral green or to improve the saturation of the three-colour matching field by the addition of the complementary to green. Thus, to match the spectral colours as a whole, either there must be three controls in the matching field and

Continued at foot of page 36

DEFECTIVE COLOUR VISION

W. D. WRIGHT A.R.C.S. D.Sc.

Professor of Technical Optics
Imperial College of Science and Technology
London

1 Colorimetric data
2 Classification of colour defects
3 Incidence of defective colour vision
4 Individual differences of colour vision
5 Limitations of normal colour vision
6 Acquired colour defects
7 Practical colour vision tests
8 Social consequences of defective colour vision
References

Scientific discussions of colour tend to run into difficulties because of the elusive nature of colour itself. Basically, colour should be regarded as a quality of the visual sensation, and as such is subjective and incapable of physical measurement. The quality of redness, for example, is only known by introspection and it could be, and sometimes is, argued that even amongst a group of normal observers it is impossible to be dogmatic and claim that redness has the same appearance to every member of the group. To define red as, say, the colour of blood is merely to argue in a circle, since this simply reverses the information we were given as children that blood is red.

In spite of these subjective difficulties, and they are very real, it has proved possible to make very significant quantitative studies of the mechanism of colour perception without reference to the actual appearance of colours, and many very significant data on the characteristics of normal and defective colour vision have been recorded.

1. Colorimetric Data

The ability to distinguish one colour from another implies that the retina responds differently to light of different spectral quality, and the recording of the smallest wavelength difference $\Delta\lambda$ that can be distinguished by a change of hue is an observation that can be made without any description of the hue itself; and when the observation is repeated at different wavelengths λ through the spectrum, the wavelength discrimination curve obtained by plotting $\Delta\lambda$ against λ is directly relevant to an understanding of the retinal processes responsible for colour perception. Indeed, this is perhaps the most significant curve we have in separating normal from defective colour vision.

We can also compare the relative brightness (or luminance, to use a new photometric term introduced to avoid certain ambiguities associated with brightness) of different colours in the spectrum. We can, for example, record the relative energy e_λ at wavelength λ required to produce a given sensation of luminosity and by plotting $\frac{1}{e_\lambda}$ against λ we obtain the luminosity curve of the eye (or the curve of relative luminous efficiency of monochromatic radiation). For brevity, we will call this curve the V_λ curve, since V_λ is the symbol given to the ordinates of the curve when this is adjusted to have a maximum value of unity. The V_λ curve is of physiological importance, since it is a measure of the spectral sensitivity of

---

CHROMATIC ABERRATION AND THE POLYCHROMATIC THEORY
Continued from page 35

three additional controls in the matched field, or there must be six controls in the matching field, one for each of the three matching colours, and one for each of the complementaries to those matching colours. In either case a total of six controls is necessary.

6. Summary

The fact is firmly established that the lens system of the eye suffers from chromatic aberration. Why then does correcting this aberration by means of a Helmholtz lens produce a negligible effect on visual acuity, and why do not the coloured fringes produced by this aberration clearly show themselves when the eye is being used for normal vision? The reply is that there is a correcting mechanism of a special kind which can not only eliminate the fringes produced by the lens system of the eye but also those produced by lenses which are being used in conjunction with the eye, such as spectacle lenses and hand magnifiers. This mechanism eliminates yellow and at the same time lowers its intensity without similarly affecting red and green. It is evident that yellow has a separate path to the brain. This mechanism is not limited to the fovea but operates all over the useful retina. The polychromatic theory is compatible with this mechanism, and with the trichromacy and other properties of human colour vision.

REFERENCES

Abney, W. & Festing, E. R. (1892) *Philos. Trans.* **183,** 537
Dartnall, H. J. A. (1950) *Nature, Lond.* **166,** 207
Dollond, J. (1753) *Philos. Trans.* **48,** 289
Euler, — (1765) *J. Encyclop.* **2,** 146
Göthlin, G. F. (1943) *The fundamental colour sensations in man's colour sense*, p. 70. Almqvist & Wiksell, Stockholm
Granit, R. (1947) *The sensory mechanisms of the retina: with an appendix on electroretinography.* Cumberlege, London
Hartridge, H. (1947) *Philos. Trans.* B, **232,** 510
Hartridge, H. (1950) *Recent advances in the physiology of vision.* Churchill, London
Helmholtz, H. von (1909) *Handbuch der physiologischen Optik*, 3rd ed., p. 146. Vose, Leipzig
Ivanoff, A. (1947) *Rev. Opt. (théor. instrum.)* **26,** 145
König, A. (1894) *S. B. Akad. Wiss. Berl.*, p. 577
Luckiesh, M. & Moss, F. K. (1937) *The science of seeing*, p. 422. Van Nostrand, New York
Purdy, D. M. (1931) *Amer. J. Psychol.* **43,** 541
Shlaer, S., Smith, E. L. & Chase, A. M. (1942) *J. gen. Physiol.* **25,** 553
Sloan, L. L. (1928) *Psychol. Monogr.* **38,** No. 1
Thomson, L. C. (1952) *Ophthal. Lit.* **6,** 3
Walls, G. L. & Mathews, R. W. (1952) *Univ. Calif. Publ. Psychol.* **7,** 1
Willmer, E. N. & Wright, W. D. (1945) *Nature, Lond.* **156,** 119
Wright, W. D. (1946) *Researches on normal and defective colour vision.* Kimpton, London

the eye, and again significant differences are found between the normal observer and some types of colour defect.

A third important characteristic concerns colour mixture and colour-matching. Experiment shows that the normal observer can match colours by the additive mixture of three suitably chosen radiations—saturated red, green and blue stimuli being most commonly employed. This has provided the chief experimental basis for the three-receptor theory of colour vision, in which it is assumed that there are three types of receptor in the retina, one primarily red-sensitive, a second green-sensitive and the third blue-sensitive. However that may be, it is possible, by making a series of tristimulus colour-matches of the colours in the spectrum, to break up the V_λ curve into three component curves which would appear to be closely related to the spectral sensitivity curves of the receptor processes postulated by the three-receptor theory. Again, these curves are of the greatest interest in defective colour vision, since in certain types of vision—to which the general term dichromatism is applied—colour-matching requires the additive mixture of only two stimuli, and the V_λ curve can be split into two components. This inevitably suggests that the defect is due to the existence of only two types of receptor instead of three, with different pairs of receptors being active in different forms of the defect. It is possible to explain on this basis why in some types of defect colour discrimination may be weakest in the yellow part of the spectrum, while in another type of defect the observer cannot discriminate the blue-green wavelengths.

None of these experimental data requires that in their determination the subjective appearance of the colours should be named, and their interpretation is free from the uncertainties which cannot be avoided in tests in which naming is required. Indeed, it is very reasonable to suppose that psychophysical data of this kind can be fairly directly interpreted in terms of the retinal processes, and perhaps for this reason the data provide a surprisingly neat classification of the different types of colour defect. There is little reason to doubt that if all colour defectives could be adequately tested on the rather elaborate colorimetric equipment needed to measure the colour mixture and colour discrimination curves, then the vast majority of them could be fitted into one or other of the orthodox groups. Occasionally, of course, we should have to expect some unusual feature in a subject's colour vision that did not readily fit into the classification, for in the elaborate pathway connecting the initial light-sensitive reaction in the eye with the final development of a colour sensation in consciousness, there is ample opportunity for some uncommon abnormality to develop. This would in no sense, however, invalidate the basic classification.

2. Classification of Colour Defects

Observers may be divided into three main groups according to the number of knobs they have to use in an additive colorimeter when making colour-matches. The names customarily used to designate the three types, and the number of controls needed by them when they use a colorimeter, are as follows:

I Monochromat—one control
II Dichromat—two controls
III Trichromat—three controls

Whether better names could have been chosen is a matter for debate, but for historical reasons they can hardly be changed now. Confusion may arise for those with little knowledge of the subject, since it is sometimes believed that the trichromat can perceive only three distinct colour sensations, the dichromat two and the monochromat one. This is an assumption for which there is no justification, and even in the simplest case of the monochromat—who has no colour discrimination—his perception ranges between the two sensations of white and black.

Each of the main groups can be subdivided further. Two forms of monochromat can be distinguished, according to their V_λ curves. In one case the maximum of the curve occurs at a wavelength around 0.51 μ. and in the other at about 0.56 μ. Measurements show that the two curves correspond very closely to the rod and cone spectral sensitivities respectively and other evidence supports the idea that the two forms are due to either pure rod or pure cone activity. We may therefore subdivide Group I into:

I (i) Rod monochromat
(ii) Cone monochromat

Both types are very rare, especially the latter, and rod monochromatism is generally associated with other pathological defects of vision such as photophobia and nystagmus.

Three types of dichromat have been studied in considerable detail and their existence is well established. They can be distinguished by the parts of the spectrum where colour discrimination is extremely poor, by the choice of the two stimuli that have to be used in colour-matching and by their V_λ curves. The names commonly used to distinguish the three types are:

II (i) Protanope
(ii) Deuteranope
(iii) Tritanope

In protanopia and deuteranopia, there is little or no colour discrimination in the red-yellow-green range of the spectrum, and for this reason these types are often referred to as red-green defectives, but in tritanopia it is in the green-to-blue wavelengths that discrimination is lacking. In protanopia and deuteranopia the stimuli used in colour-matching include one from the blue end of the spectrum and the other from the red-yellow-green range; in tritanopia, one is chosen from the red and the other from the green-to-blue range. In deuteranopia and tritanopia, the V_λ curve differs little from the normal photopic curve and has its maximum between 0.55 μ. and 0.56 μ.; in protanopia the maximum occurs at about 0.54 μ. and the ordinates at the red end of the spectrum are very much lower than the normal.

These characteristics are consistent with the idea that protanopia is due to the inactivity of the red-sensitive receptors, deuteranopia to the fusion of the red and green activities, and tritanopia to the inactivity of the blue receptors or possibly to the fusion of the green and blue processes. Those who do not accept the three-receptor theory will no doubt have other explanations, but this interpretation at least helps to provide a mental picture of the nature of the defects. The most exhaustive studies of protanopia and deuteranopia have been reported by Pitt (1935) and that of tritanopia by Wright (1952). It is theoretically possible that other types of dichromat might occur, but these have not so far been well established.

The trichromats may be subdivided into the normal trichromats and the anomalous trichromats. The former have good colour discrimination and their colour-matches are

similar to the average matches for the majority of the population. The anomalous trichromats have poorer colour discrimination than the normal and although three stimuli have to be used in colour-matching, the precision of setting is lower than in normal colour-matching, especially with certain ranges of colours, and the matches themselves may differ from the normal matches. The anomalous trichromat appears to have a form of colour perception intermediate between the normal and the dichromat, so that the trichromat classification can be subdivided as follows:

III (i) Protanomalous trichromat
(ii) Deuteranomalous trichromat
(iii) Tritanomalous trichromat
(iv) Normal trichromat

The most exhaustive study of protanomalous vision has been reported by McKeon & Wright (1940) and of deuteranomalous vision by Nelson (1938). No detailed study of tritanomalous vision has been made as yet.

3. Incidence of Defective Colour Vision

Various large-scale surveys have been made from time to time by different workers using a variety of tests, and from a recent summary by Judd (1943) we can extract the values given in Table I.

TABLE I. INCIDENCE OF DIFFERENT TYPES OF COLOUR VISION

Group	Percentage of population	
	Male	Female
I Monochromatism . .	Very rare	Very rare
II Dichromatism:		
Protanopia . .	1.0	0.02
Deuteranopia . .	1.1	0.01
Tritanopia . .	0.005*	0.003*
III Trichromatism:		
Protanomaly . .	1.0	0.02
Deuteranomaly . .	4.9	0.38
Tritanomaly . .	Fairly rare, but not known accurately	Fairly rare, but not known accurately
Normal trichromat .	92.0	99.6

* Figures for tritanopia have been taken from a recent investigation by Wright (1952).

4. Individual Differences of Colour Vision

The implication in any classification such as that given above is that persons falling in any given section have closely similar vision. While this is true in essence, it would be grossly misleading to deny the existence of significant differences within each group. Pigmentation, for example, in both the optic media and the retina is one obvious source of individual variation affecting both the V_λ curve and the colour mixture curves. It cannot affect the wavelength discrimination curve, but it can affect discrimination tests using dyed or painted surfaces, or colour filters, since the spectral composition of the light reaching the retina will be changed by absorption in the eye. Again, while a dichromat might have only two kinds of receptor in his retina, it is easy to visualize that their relative sensitivities might differ from one observer to another, perhaps through differences in photochemical concentration, perhaps through differences in relative numbers of receptors, perhaps through differences of facility of transmission along the nerve paths, or from some other cause of this kind. Whether the spectral sensitivities of individual receptor processes could vary is problematical; in so far as the sensitivity curve is determined by the spectral absorption curve of some fixed photochemical compound, then any such variation would seem unlikely, but the possibility of inactive substances acting as internal filters within the receptors cannot be excluded. However that may be, some variation within each section of the classification can be accepted without requiring any basic modification in the classification itself.

Nevertheless, in the case of anomalous trichromatism the situation is different. It is widely recognized that varying degrees of this defect can exist, extending from the nearly normal to the nearly dichromat; this spread was apparent, for example, in the reports by Nelson (1938), McKeon & Wright (1940) and Wright (1946). Since the origin of anomalous trichromatism is by no means clear, there is no occasion here to speculate on how these varying degrees of defect can arise, but the evidence of the wavelength discrimination curve is sufficient to show that there is a genuine variation in colour discrimination which cannot be attributed, for example, to differences of pigmentation.

This is the seat of the major problem in practical colour vision testing, where a decision has to be made as to the fitness of persons to carry out some particular task. Somewhere a line has to be drawn between those who can be accepted and those who must be rejected, and it is extremely difficult to decide on the place where the line should be drawn. For the reasons outlined above, we find variations in colour vision among people who should still be regarded as normal, yet in one or other of their characteristics they might well overlap with anomalous trichromats whose vision is in some measure defective. Rejection for a job must obviously depend in part on the nature of the job—whether, for instance, the question of safety is involved—but it is nevertheless extremely difficult to decide whether one or two mistakes in, say, the lantern test or the Ishihara test should or should not be ignored. Opinions are divided on this, and there is no immediate prospect of any *obiter dicta* being forthcoming to remove the uncertainty. But, at least in jobs other than those requiring the correct recognition of signals, the view expressed in a Report of the Physical Society (1946) should be noted. The Committee concerned stated:

> We must, however, put on record that one of our most persistent findings is that some colour-defective observers can carry out colour operations successfully where we might have anticipated errors or failure. Except, therefore, where the operations are of an especially critical nature, or where questions of safety are involved, we advocate that doubtful cases should be decided in favour of the candidate, provided he is warned that he may be handicapped; for example, through loss of speed, and in suffering greater strain.

This question of individual differences in colour vision has been discussed at length by Pickford (1951)[1], and his book provides extensive and valuable evidence of the kind of differences that may be discovered with tests made on large groups of people. Nevertheless, the analysis and classification of the differences is rendered difficult because the testing was

[1] Reviewed on p. 79 of this Bulletin.

not carried out with apparatus using spectral stimuli. In mass testing this is generally impossible, but if such apparatus is not used, then it has to be accepted that a complete analysis is virtually impossible, because of the many different ways in which the variations in the colour vision functions may combine. An adequate classification requires the determination of the V_λ curve, the colour mixture curves and the wavelength discrimination curve. This is impracticable in routine colour vision testing, but in a scientific study it is very desirable indeed that as far as possible each variable should be studied independently.

5. Limitations of Normal Colour Vision

Under certain conditions of viewing, an observer with otherwise normal colour vision may find his colour discrimination defective and his visual characteristics tending to simulate those of the congenital colour defectives. The easiest case to demonstrate is vision at very low illuminations when colour discrimination becomes impossible. It seems quite certain that under these conditions only the rod receptors are functioning and the observer in effect suffers from rod monochromatism.

At intermediate levels of illumination when both the rod and cone mechanisms are presumably functioning together, the work by Gilbert (1950) suggests that, in parafoveal vision, colour perception tends towards the tritanopic form of colour defect. Weale (1951) also reports a tritanopic tendency at low illuminations in the extrafoveal retina at 10° and 15° from the fovea. On the other hand, Thomson & Trezona (1951) found that for foveal vision with a 1° 20′ field, while discrimination deteriorated as the illumination diminished, the deterioration occurred at the same rate over the greater part of the spectrum, and they did not report any tritanopic tendency.

A further example of colour deficiency in normal observers occurs with small fields. As long ago as 1894, König reported that when fields subtending only a small angle at the eye (up to some 20 minutes of arc) were viewed at the fovea, a colour defect was exhibited corresponding to tritanopia. The ability to discriminate greens and blues virtually disappeared, and yellows could be matched by a mixture of red and blue stimuli. The effect is difficult to observe, because of the need to maintain accurate fixation on the small field, but quantitative confirmatory experiments have been reported by Willmer (1944), Willmer & Wright (1945) and Thomson & Wright (1947). While the central fovea, compared with neighbouring zones of the retina, has been shown by Stiles (1949) and by Thomson & Wright (1947) to possess a marked lack of sensitivity to blue light, yet, as Hartridge (1945) has shown, other points on the retina have a tritanopic tendency under small-field conditions. However, considering the generally poorer discriminatory power of the retina away from the visual axis, the tendency is relatively less significant than in the central fovea.

Small-field tritanopia has important practical consequences, particularly in the choice of signal colours. Blue and green, for example, cannot be used in the same signalling system, neither should yellow and white, and it was interesting to find that in the course of an investigation on the recognition of signal colours, Holmes (1941) reported an island of colours on the purple side of the chromaticity chart which were liable to be named as yellow. This was a confusion quite characteristic of small-field tritanopia.

6. Acquired Colour Defects

This article is primarily concerned with the congenital form of defective colour vision, but reference should be made to the fact that the defect can be acquired through certain toxic conditions and may develop as a result of certain diseases. Information on this aspect has been summarized by Judd (1943), but the relationship between the acquired and the congenital defects has not yet been clearly established, owing to the absence of detailed quantitative studies on the acquired defect. There is evidently scope for research on the V_λ curve, and on the colour mixture and the wavelength discrimination curves, in selected cases of acquired defective colour perception.

7. Practical Colour Vision Tests

The usual reason for carrying out colour vision testing is to grade a person's vision for a particular task. For this purpose it is sound in principle to use a test which as nearly as possible simulates the task itself, but it is comparatively rarely that the task is of a sufficiently limited and clearly defined nature that the principle can be put into practice effectively. In the lantern test, however, we do have a well-established and widely used test which does set out to imitate the task of recognizing signal lights, and it is customary to use this test for grading seamen, engine-drivers, pilots, etc. In other cases a test not so closely tied to a particular application may be adequate, for example, the Ishihara test in which inability to discriminate colours may lead to mistakes in reading the test cards.

Ideally, the perfect test would give a clear-cut indication whether a subject has normal or defective colour vision and might give additional information about the type of defect. For reasons which have already been discussed, in border-line cases we are unlikely to reach the straightforward decision we should like; moreover, in addition to the colour vision functions varying in different ways, other factors may also tend to confuse the issue. The temperament of the observer, for example, is one: whether he is nervous, or habitually finds difficulty in making up his mind; whether he is capable of concentrating or is liable to let his attention wander, and so on. Or again, the observer may have some refractive error which may add to his difficulty in reading a chart or focusing on a small illuminated aperture. There is also the technique of the operator to be considered; one operator may put one type of question to the candidate, another may ask rather different questions or use a different tone of voice. There should not be much difficulty in laying down clearly defined procedures for the operator to follow, but it is perhaps not really desirable to make colour vision testing too mechanical and robot-like. Then there is the test itself: one model may not be identical with another; the quality of the lighting may not be quite correct; the intensity of the lantern lights may be too high or too low; the colour printing of the confusion charts may be slightly degraded, and so on.

All these points can, of course, be attended to, but at a price. The tests themselves may for economic reasons have to be kept down to a reasonable cost, and the amount of time available to attend to the more refined points of the test may be very limited in mass testing. Indeed, an important feature of a test may be the speed at which it can be carried through. We are forced, therefore, to the conclusion that while there is little difficulty in detecting a person with a gross colour defect very quickly, the border-line cases are likely to remain a problem for some time to come.

With regard to the various tests available, a Report of the Physical Society (1946) gives an excellent summary of them and of their potentialities. Reports have also been published in the *Journal of the Optical Society of America* of tests made to compare the different editions and the different types of confusion chart test (Chapanis, 1948, 1949; Hardy, Rand & Rittler, 1945). The difficulty then arises that with each new edition of a test a new assessment is necessary. Thus a tenth Japanese edition of the Ishihara charts[2] has been issued (Ishihara, 1951), which is available in England through H. K. Lewis & Co. Ltd. and, while this edition will no doubt be successful in detecting colour defectives, some of the printing is undoubtedly darker and more saturated than in earlier editions; for this reason, border-line cases at least are likely to make different mistakes and perhaps a different number of mistakes according to which edition is used. It is too early to say whether the new edition will be more satisfactory than the older editions; the point that has to be made is that differences exist between them.

It follows that, however desirable it may be to standardize the tests, at the present time the technique of testing and the interpretation of the results of the tests must to some extent be left to the discretion of the examiner. The tools he chooses will inevitably depend on the purpose for which the testing is being carried out, but a typical battery of tests would include a confusion chart test such as the Ishihara test or the Swedish Böstrom-Kugelberg test, a Nagel-type anomaloscope (now available from Schmidt and Haensch in Germany), a lantern test such as the Martin lantern (Martin, 1939, 1943) or one in which equal care had been taken in the standardization of the components, and the Farnsworth-Munsell 100-hue and dichotomous test (Farnsworth, 1943). Where appropriate, suitable trade tests would also be included. Only when a detailed knowledge of the colour vision functions was required would the more elaborate colorimetric apparatus (Wright, 1946) have to be included.

8. Social Consequences of Defective Colour Vision

The practical effect of defective colour vision is essentially that the number of different colours that can be distinguished by the colour defective is significantly less than the number seen by the normal person. One of the consequences of this is that his colour vocabulary includes a number of—to him—redundant terms and he is liable, especially in certain colour ranges, to use the wrong name to describe a colour. Hence when he is in communication with a normal person, the possibility of ambiguity and misunderstanding is always present.

Not only, however, may he misname a colour, but he may act incorrectly. The best-known example of this is the incorrect recognition of a signal light, but its importance is now so widely understood that there is no occasion to enlarge further on its implications here. All that does need to be stressed is that when boys are planning to go into one of those occupations involving signal recognition, they should be advised at as early an age as possible to have an adequate colour vision test. It can still happen that boys may undergo several months, or even a year or more, of preliminary training for a job to which they are subsequently refused admission on account of their colour defect.

There are, in addition, many other occupations calling for good colour perception, and it is often desirable that entry to them should be preceded by a colour vision test. These occupations, which are discussed in considerable detail in the Physical Society Report (1946), include not only the more obvious colour industries such as the dyeing, textiles, paint and printing industries, but also certain branches of electrical engineering, photography, oil refining, chemistry, geology, medicine, agriculture, and so on.

In agriculture, colour is one of the most important characteristics by which the quality of a product may be judged, and the judgement may have to be made not only during the growth of the plant, but at various stages in marketing. Many examples could be quoted both of the economic importance of correct colour judgements and of personal difficulties when persons with colour defects become involved in the judgements. Typical produce in which this can occur includes tomatoes, tobacco, cotton, apples and bananas.

In electrical engineering, as in a number of other industries, colour discrimination is required in order to identify the colour codes which are widely used in the electrical wiring of radio and telephone equipment, and the like. Sometimes the areas of colour are extremely small, as in the marking of resistors, and when the parts are dirty, even a person with normal colour vision finds it hard to see the colours correctly. Wherever possible, colour codes should be ancillary to pattern codes, and the colours which are chosen should generally be as saturated as possible. Apparently minor considerations, such as the advantage of glossy over mat paint, may be important in assisting the colour defectives.

In geology and geography, colour codes on maps and drawings are very common, but too often the colour key is located at a remote corner of the map, and the areas of colour shown in the key are too small. Both these conditions militate against their easy identification by the colour defective, and the frequent choice of colours from the green-brown range is the worst possible from their point of view.

These examples are sufficient to indicate that a little thought and knowledge could often make the colour defective's task much easier than it usually is. Also, if a person is known to be colour defective, then it may be quite easy to arrange that he is not given tasks which are beyond his capacity. There are some occupations which a person with defective colour vision should be strongly discouraged from following, otherwise he may find himself at a serious disadvantage as time goes on; he may, for instance, be debarred from promotion into a higher grade of work requiring more critical colour discrimination or greater responsibility for correct decisions about colour.

Many examples could be given of cases where great hardship has been caused through the diagnosis of a colour defect at too late a stage in a person's career, so that anyone concerned with vocational guidance should make himself familiar with the difficulties that can arise. In particular, it is greatly to be hoped that ways and means will be found to include the testing of colour vision as a regular item in school medical inspections.

[2] Reviewed on p. 82 of this Bulletin.

REFERENCES

Chapanis, A. (1948) *J. opt. Soc. Amer.* **38**, 626
Chapanis, A. (1949) *J. opt. Soc. Amer.* **39**, 242
Farnsworth, D. (1943) *J. opt. Soc. Amer.* **33**, 568
Gilbert, M. (1950) *Proc. phys. Soc. Lond.* B, **63**, 83

Continued at foot of page 41

VISUAL PROPERTIES STUDIED BY SUBJECTIVE MEASUREMENTS ON THE COLOUR-ADAPTED EYE

W. S. STILES O.B.E. D.Sc.

National Physical Laboratory
Teddington, Middlesex

1 Subjective methods of investigating colour adaptation
2 Visual mechanisms: spectral sensitivity
3 Monocular colour-matching after conditioning
4 Results from binocular colour-matching
5 Separation of the rod and cone mechanisms by the two-colour threshold technique
6 Threshold analysis of the cone mechanism
7 Relation of the threshold and colour-matching mechanisms
References

1. Subjective Methods of Investigating Colour Adaptation

One way of probing the visual mechanism in order to discover its properties is to expose the eye to coloured light of various intensities and to make subjective observations while the mechanism is still in the perturbed state brought about by the coloured conditioning stimulus. In recent years this principle has been followed in two distinct types of experiment: binocular colour-matching (Wright, 1934, 1937) and the two-colour threshold technique (Roaf, 1932; Stiles & Crawford, 1933; Stiles, 1939). In the former, one eye, say the right, is exposed to the conditioning stimulus—which is commonly an extended field of uniform brightness and colour—, the other eye being kept in the dark. After the conditioning stimulus has been removed, the apparent colour of a small test stimulus (monochromatic as a rule), presented to the right eye only, is measured. This is done by matching it with a comparison stimulus of similar size consisting of a mixture in adjustable proportions of three differently coloured primary stimuli, which is presented to the left eye only. The quantities of the three primary stimuli required to match the test stimulus provide a measure of its apparent colour in the same trichromatic system as is used for specifying colours in the normal way. The effect of the conditioning stimulus is always to depress the sensitivity of the eye so that, in the first measurement after it has been extinguished, the three quantities in the matching mixture have low values. As recovery proceeds, they rise and their values determined in succeeding measurements trace out three colour-recovery curves. But in general the three quantities rise at different rates and there is usually a progressive change in their relative values—which determine the resultant colour—as the effects of the conditioning stimulus pass off.

The experimental procedure in the two-colour technique is quite different. No use is made of colour-matching, the quantity measured being the least perceptible intensity of a small test stimulus—usually monochromatic—which is exposed as a flash at a particular part of the visual field. As before, the conditioning stimulus is commonly an extended area of uniform brightness but in general it is exposed continuously and is not extinguished before the application of the test stimulus (but see section 7). Thus the retinal area on which the image of the test stimulus is formed receives, during the test spot exposure, an illumination both from the test and the conditioning stimulus. The test stimulus is in this sense an addition or increment and the intensity at which it can just be detected—whether as an increase of brightness or as a change of colour is not distinguished—is termed the increment threshold. An important feature of the method is the use in general of different colours for the test and conditioning stimuli. Whereas, in the binocular matching method, the test stimulus is exposed at a time when the visual mechanism is certainly not in a steady state—it is in process of re-adapting to darkness—, in the two-colour technique the observation is normally made when the eye is in equilibrium with the conditioning stimulus. Unlike the mixture quantities in colour-matching experiments, the increment threshold has no obvious relation to colour sensations; but by its aid it is possible to determine the spectral sensitivities and some other properties of visual mechanisms which are almost certainly concerned in colour perception.

2. Visual Mechanisms: Spectral Sensitivity

The starting-point in the interpretation of the results of both of the above methods is the assumption that in any particular region of the visual field—or in the corresponding region of the retina as it is more convenient to put it—vision is served not by one but by several mechanisms, these mechanisms having different spectral sensitivities and being responsible for the variety of colour sensations experienced. No precise definition of such a component mechanism can be given. It may be supposed that in the retinal region considered some of the rod and cone end-organs—a selected

DEFECTIVE COLOUR VISION
Continued from page 40
Hardy, L. H., Rand, G. & Rittler, M. C. (1945) *J. opt. Soc. Amer.* **35,** 268; 350; 481
Hartridge, H. (1945) *Nature, Lond.* **155,** 657
Holmes, J. G. (1941) *Trans. Illum. Engng Soc.* **6,** 71
Ishihara, S. (1951) *Tests for colour-blindness,* 10th ed. Lewis, London
Judd, D. B. (1943) *J. opt. Soc. Amer.* **33,** 294
König, A. (1894) *S. B. Akad. Wiss. Berl.,* p. 577
McKeon, W. M. & Wright, W. D. (1940) *Proc. phys. Soc. Lond.* **52,** 464
Martin, L. C. (1939) *Brit. J. Ophthal.* **23,** 1
Martin, L. C. (1943) *Brit. J. Ophthal.* **27,** 255
Nelson, J. H. (1938) *Proc. phys. Soc. Lond.* **50,** 661
Physical Society. A Committee of the Colour Group (1946) *Report on defective colour vision in industry.* Physical Society, London
Pickford, R. W. (1951) *Individual differences in colour vision.* Routledge & Kegan Paul, London
Pitt, F. H. G. (1935) *Spec. Rep. Ser. med. Res. Coun., Lond.* No. 200
Stiles, W. S. (1949) *Docum. ophthal.* **3,** 138
Thomson, L. C. & Trezona, P. W. (1951) *J. Physiol.* **114,** 98
Thomson, L. C. & Wright, W. D. (1947) *J. Physiol.* **105,** 316
Weale, R. A. (1951) *J. Physiol.* **113,** 115
Willmer, E. N. (1944) *Nature, Lond.* **153,** 774
Willmer, E. N. & Wright, W. D. (1945) *Nature, Lond.* **156,** 119
Wright, W. D. (1946) *Researches on normal and defective colour vision.* Kimpton, London
Wright, W. D. (1952) *J. opt. Soc. Amer.* **42,** 509

proportion—are associated together by the organization of the nerve paths along which the nerve pulses initiated by light in these end-organs are conducted via various synapses in the retina, geniculate bodies and visual cortex to regions of the brain from which the efferent pulses of the response—oral descriptions, hand signals, manipulations of colorimeter controls and the like—are set going. In general the end-organs of a particular mechanism would be disseminated in the retina among end-organs of other mechanisms. In one extreme case, the end-organs might all be of identical type and properties and each might be equipped with its own exclusive retinal bipolar cell and retinal ganglion cell, so that in the retina the paths from the end-organs would be quite independent, and the association by which they could be said to belong to one mechanism would necessarily be located at higher levels. The kind of central association here in mind is sometimes expressed, in the lack of any actual knowledge of the processes called into play in the brain, by saying that the paths lead to a particular centre—the " red " centre, let us say, if the mechanism in question is supposed to be one which when stimulated leads to the verbal response " red " from the subject. Studies of the neural organization of the retina itself (the primate retina), particularly the work of Polyak (1941), make it clear that exclusive paths from end-organ to optic nerve fibre, although they probably occur to some extent in the central fovea, are not typical. Not only are the end-organs grouped by connecting with common bipolar cells, and bipolars by connecting with common ganglion cells, but horizontal and so-called amacrine cells provide numerous lateral connexions, among the end-organs in the first case, and among the bipolars and ganglion cells in the second. Moreover, outside the rod-free area, rods and cones are often linked to common bipolar cells, and the same end-organ may participate in more than one neural grouping. It seems probable therefore that a mechanism may well include end-organs of different type and that complete independence of mechanisms even at retinal levels may be exceptional.

Both an end-organ and a visual mechanism may possess one or more relative spectral sensitivity curves, in the sense that, when they are stimulated in turn by monochromatic radiations of different wavelengths, the same effect or response may be produced, provided that the intensity of the stimulus is suitably adjusted. Such a spectral sensitivity curve is defined as the reciprocal of the necessary monochromatic stimulus intensity (expressed in energy units or, better, in terms of the number of light quanta) plotted against the wavelength or frequency of the light. Although it has not yet proved possible to tap into the visual mechanism to observe directly the effect of light stimuli on the end-organs (at least in the higher animals), it is probable that, whatever the character or the magnitude of the effect which the monochromatic stimuli might be required to produce, the relative intensities for different wavelengths would be the same, i.e., the end-organ would have a unique relative spectral sensitivity curve. This assumes that the end-organ contains only one photosensitive substance, that its " concentration " is not too great for the law of absorption by dilute solutions to apply, and that any selectively absorbing material (filter material) through which the light has to pass before reaching the photosensitive substance remains unchanged under various stimulus conditions. In those circumstances the change with wavelength of the photosensitivity of the photosensitive substance and the spectral transmission of any permanent filter material determine the relative spectral sensitivity O_λ of the end-organ. The effect of a monochromatic stimulus of energy E_λ depends solely on the product $E_\lambda O_\lambda$, or, for a stimulus containing energy in several wavelengths, on the sum $\Sigma E_\lambda O_\lambda$ (linear-additive law). The retinal rods seem to possess this straightforward kind of spectral sensitivity. A visual mechanism equipped with a single type of end-organ satisfying the conditions enumerated above would also have a unique relative spectral sensitivity curve—the same as that of its end-organs—and would conform to the linear-additive law. But if two or more end-organs were associated together in the mechanism, the spectral sensitivity M_λ of the latter would certainly be some kind of composite of the spectral sensitivities A_λ, B_λ, etc., of its end-organs, but not necessarily a simple one. It would obviously depend on the way nerve pulses from the different types of end-organ combined their effects in firing higher order neurones of the mechanism. The simplest combination, mathematically, would for M_λ to be linearly dependent on A_λ, B_λ, etc.: $M_\lambda = \alpha A_\lambda + \beta B_\lambda$ where α and β are constants, but plausible physiological models for this and other simple laws of combination are difficult to obtain. Two questions of importance in colour-adaptation studies suggest themselves: will the spectral sensitivity of a compound mechanism (i.e., one involving more than one type of end-organ) conform to the linear-additive law, and will the spectral sensitivity curve remain stable under differently coloured conditioning stimuli which might be expected to depress to different extents the sensitivities of end-organs of different types? These questions and the possibility of the end-organs themselves deviating from the conditions specified earlier must be borne in mind in working with the notion of a visual mechanism.

3. Monocular Colour-Matching after Conditioning

If a test stimulus and a comparison mixture of three primary stimuli, viewed together by one eye, are made to match, and the eye is then subjected to a white or coloured conditioning stimulus of moderate intensity, the test and comparison stimuli on being re-viewed still appear to match although their common colour may be considerably changed. This result, together with the ordinary laws of trichromatic colour-matching, is consistent with the view that in the foveal retina there are just three types of end-organ, each with a spectral sensitivity of linear-additive character. It is immaterial whether the visual mechanisms concerned are simple or compound, and whether for compound mechanisms their spectral sensitivities are linear-additive and stable. But more complicated alternative views are also possible. For example, there could be four types of end-organ, two of which only occur associated in a compound mechanism having a linear-additive and stable spectral sensitivity, so that there would still be only three spectral sensitivity curves in play. These curves are not uniquely determined by the colour-matching data but the latter yield three experimental curves—x_λ, y_λ, z_λ, let us say—in terms of which the spectral sensitivities of the three end-organs (or stable mechanisms according to the more complicated view) are linear combinations such as $ax_\lambda + by_\lambda + cz_\lambda$, the three sets of coefficients a, b, c being undetermined.

It was shown by Wright (1936) that, if sufficiently intense conditioning stimuli are used, monocular colour-matches viewed after exposure no longer hold good. Conditioning

stimuli of some 15,000 trolands[1] or more are required (15,000 trolands corresponds to a brightness between those of the clear sky and the ground at noon in summer). The figures given in Table I for one experiment show how the amounts of red (wavelength 650 mμ.) and green (wavelength 530 mμ.) primaries in a mixture matching a yellow (wavelength 582.5 mμ.) test stimulus are modified after a conditioning stimulus of 48,000 trolands of white light.

TABLE I. THE EFFECT OF INTENSE WHITE CONDITIONING ON A MONOCULAR COLOUR-MATCH (Wright, 1936)

State of the eye	Amounts of the primaries in a mixture matching yellow		Ratio: amount of red / amount of green
	Red primary	Green primary	
Before conditioning	836	552	1.51
After conditioning	1,189	445	2.67

The effect is clearly considerable, and similar observations on other monochromatic and on white light test stimuli also gave definite changes. By an ingenious method of modifying the units of the primaries before and after conditioning, Wright was able to test whether the colour changes could be attributed wholly to the production or modification of selective filter material affecting the light reaching all the end-organs. He found that this was not possible, so that the colour changes must be the result, in part at least, of filter material affecting the end-organs unequally or of a change in the spectral photosensitivities of the actual photosensitive substances. If a compound mechanism, incorporating an extra type of end-organ and with a normally stable spectral sensitivity, were present in the retina, a break-down of this stability under intense conditioning might be responsible for the effects. But if there are only three types of end-organ no possible neural disorganization which leaves the relative spectral sensitivities of the end-organs intact would provide an explanation. Unfortunately only scanty observations are as yet available on this significant deviation from the usual trichromatic conception.

4. Results from Binocular Colour-Matching

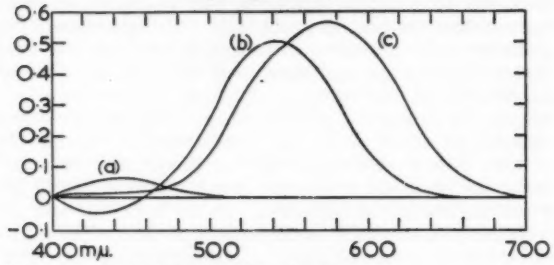
The possibility of monocular colour-matching with a mixture of three primaries does not necessarily imply that the same thing can be done in binocular colour-matching when the two eyes are in very different conditions. More than three visual mechanisms might have to be responding similarly to test and comparison stimuli for a match. This creates no difficulty in the monocular case because the three types of end-organ feeding all the mechanisms are in the same condition for the two stimuli. Provided the latter are equivalent for the end-organs, all the mechanisms must be similarly activated. This argument does not apply for binocular colour-matching. However, by introspection it is judged that there are only three qualities to be equated in colour-matching—brightness, hue and saturation—, and in fact satisfactory binocular colour-matches can be made with only three primaries. It appears therefore that only three mechanisms, independent in the mathematical sense, are involved.

[1] The troland (or photon) is the photometric unit of retinal stimulation: one troland corresponds to a brightness (luminance) of one candle per square metre seen through a pupil of area one square millimetre.

But one limitation may be noted. Even in the monocular case, with any set of three actual primaries there are always some spectral colours too saturated to match. These are dealt with by adding a little white light to the test stimulus and matching the mixture so obtained. The quantities of the primaries which match the white light by itself are then subtracted from the corresponding quantities required to match the mixture, to give the quantities—one of them negative—appropriate to the test stimulus alone. The need to admit negative quantities in this way arises from the overlapping of the spectral sensitivities of the end-organs, but their interpretation in the monocular case is clear-cut only if the linear-additive law is applicable. In the binocular case the corresponding condition is the validity of the coefficient law (see below) which is now known not to be generally true. Possibly, therefore, overlapping of end-organ spectral sensitivities is not the only factor making some test colours unmatchable with an all-positive mixture of three fixed primaries.

Turning now to the colour-recovery curves obtained by the binocular method, these were found to rise regularly—often linearly for the first 100 sec.—flattening out finally as recovery approached completion. At very high conditioning intensities, however, above about 10,000 trolands, irregularities appear. For example, for a yellow test stimulus (580 mμ.) viewed after intense white conditioning, the red primary recovery curve rose initially, became constant for about one minute and then increased regularly to its final value. Wright (1946) connects these irregularities with the break-down of monocular colour-matches under similar intense conditioning. In what follows, attention is confined to results obtained with conditioning stimuli of moderate intensity which do not affect monocular matches. Perhaps the most interesting of these have centred round the attempt to determine completely the three basic spectral sensitivities for which monocular matching data provide only expressions involving unknown coefficients *a*, *b*, *c* (see section 3). Wright (1934) devised and applied a method for doing this. Starting from the simplest assumption that in binocular colour-matching the responses of three completely independent mechanisms are being equated, each equipped with end-organs of one type or at least possessing linear-additive and stable relative spectral sensitivity curves, he argued as follows. Suppose a test stimulus can be found which stimulates only one mechanism. Whatever the character of the conditioning stimulus, the only effect it can have is to alter the strength of the response in the one mechanism concerned, so that the apparent colour of the test stimulus will remain the same. Test stimuli in the extreme red and in the blue (460 mμ.) exhibit only rather small colour changes after conditioning, but for no real stimulus is the requirement satisfied rigorously. However, by measuring binocularly the colour changes produced by differently coloured conditioning stimuli on selected test colours, it is possible to deduce that certain mixtures of three real test stimuli involving a negative amount of one component (or of two) would, if they could be realized, appear always the same colour and would stimulate one mechanism only. The theory assumes that in combining the results of different binocular colour-matches a certain " law of coefficients ", to be considered immediately, can be taken as true to a first approximation. By making this assumption, Wright succeeded in obtaining a set of three such " hypothetical " test stimuli, from which with the aid of the data of monocular colour-matching he was able to derive the spectral

FIG. 1. SPECTRAL SENSITIVITIES OF THE THREE BASIC MECHANISMS OF COLOUR VISION DERIVED BY THE BINOCULAR MATCHING METHOD (Wright, 1934)



Abscissae: wavelength of the stimulus (mμ.)

Ordinates: sensitivity defined as the reciprocal of the stimulus energy required to produce a given response (units different for the three curves)

Note: Wright uses so-called luminosity units which show what are believed to be the relative contributions of the three mechanisms, (a), (b) and (c), to the sensation of brightness.

sensitivities of the basic mechanisms (fig. 1). On the assumptions made, no mechanism could have a negative spectral sensitivity, so that the dip of curve (b) below the axis, if not due to experimental error, points to some break-down of these assumptions. Otherwise the curves are placed along the spectrum at about the places to be expected from such evidence as is afforded by the study of colour defectives.

The law of coefficients may be stated as follows: if, at a given time after extinguishing the conditioning stimulus, a test stimulus A in the conditioned right eye matches a comparison stimulus a in the dark-adapted left eye, and also a stimulus B, right eye, matches a stimulus b, left eye, then the mixed stimulus $A + B$, right eye, will match the mixed stimulus $a + b$, left eye. This implies more than that the spectral sensitivities of the mechanisms are linear-additive. It implies that the factor by which the response of a particular mechanism to a given test stimulus is depressed by the conditioning stimulus does not depend on the intensity and colour of the test stimulus. Further work on Wright's method by Walters (1942), using improved apparatus, has established that considerable deviations from the law of coefficients occur. For example, on increasing the intensity of an orange test stimulus by a factor of ten, the amounts of red and green primaries required to match it after a red conditioning were increased by a factor of no more than about five. A still more striking deviation was later reported by Wright himself, as shown by Table II.

TABLE II. EXAMPLE OF A BREAK-DOWN IN THE COEFFICIENT LAW (Wright, 1946)

Test stimulus	Quantities of matching primaries immediately after removal of the conditioning stimulus (4,500 trolands, red)	
	Red	Green
650 mμ., 420 trolands . .	21	11
582.5 mμ., 200 trolands . .	2	30
650 + 582.5 (mixed) . . .	10	37

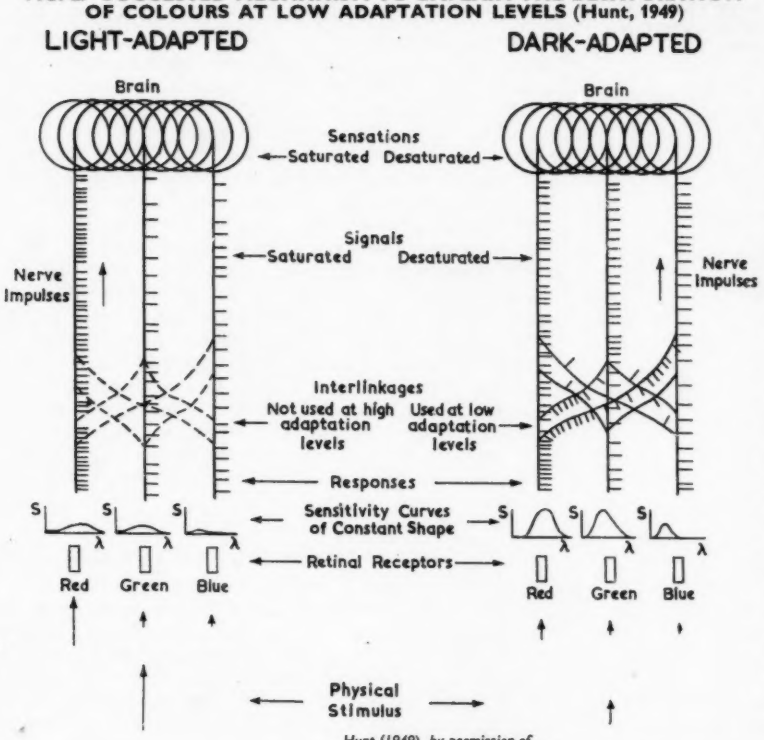
The amount of red primary required for the mixture is only about half that required for one of the constituents viewed alone. Walters found that, in general, break-down of the coefficient law made it impossible to determine the spectral sensitivities of the mechanisms in the way proposed but that he could do so by measuring colour changes for a series of test intensities and extrapolating back to the case of very low intensity. He studied only the "red" and "green" spectral sensitivities and his curves do not extend to wavelengths below about 530 mμ., but in shape and position they differ little from those of Wright.

The break-down of the coefficient law just described is attributable in part at least to the disturbing effect of the test stimulus. Other experiments by the binocular method indicate that the assumption of completely independent mechanisms is untenable. The depression of the sensitivity of one mechanism appears to depend on the extent to which the particular conditioning stimulus is affecting the other mechanisms. The interpretation of the results is difficult, but it seems that if the colour of an intense conditioning stimulus is such as to depress most the sensitivity of one mechanism, say the "red", then its effect on the "green" mechanism is less than would be expected from tests with green conditioning stimuli which have most effect on the "green" mechanism. Put in another way, a coloured conditioning stimulus seems to become more specific in its effect on the mechanisms as its intensity increases. This kind of interaction does not entail break-down of the coefficient law.

Two other noteworthy phenomena may be mentioned. Yellow conditioning stimuli produce an anomalously large depression of the sensitivity to blue test stimuli, and the initial stages of recovery with a variety of test and conditioning stimuli are marked by an unexpected need for a contribution from the blue primary in the matching mixture. This latter phenomenon of the "positive blue" violates the coefficient law, eludes explanation on the normal basis of three independent mechanisms, and is attributed by Wright (who discovered it) to "a secondary reaction occurring at a stage in the visual chain after the initial light-reception process" (Wright, 1946).

The recent work by Hunt (1949, 1950, 1952) on colour adaptation by the binocular method differs from earlier studies by the use of test stimuli extending beyond the rod-free area (they had maximum dimensions of 5° or 20°), by the removal of the very large conditioning stimulus only from the area actually required by the test stimulus, the rest being left as a surround field during the matching measurement, and, finally, by the arrangement of conditioning stimuli for both eyes. Only white ("daylight" and "artificial light") conditioning was used. The outstanding result obtained was that every test stimulus, whatever its colour, appears more saturated to the eye conditioned to the higher brightness. The hypothesis of three independent mechanisms is quite incapable of explaining this result. Hunt has proposed that while the three mechanisms may be independent when conditioned to high intensities, as the conditioning level drops, the pulses in each mechanism may tend increasingly to spill over into the other mechanisms, thereby desaturating the apparent colour of any test stimulus (see fig. 2). It must be remembered that Hunt's test stimuli fall on rods as well as cones, and he has considered whether a white sensation from the rods could cause the desaturation. This would be a possible theory if the spectral sensitivity of the rods could be represented as a linear combination of the sensitivities of

FIG. 2. SUGGESTED MECHANISM TO EXPLAIN THE DESATURATION OF COLOURS AT LOW ADAPTATION LEVELS (Hunt, 1949)



Hunt (1949), by permission of Proceedings of the Physical Society

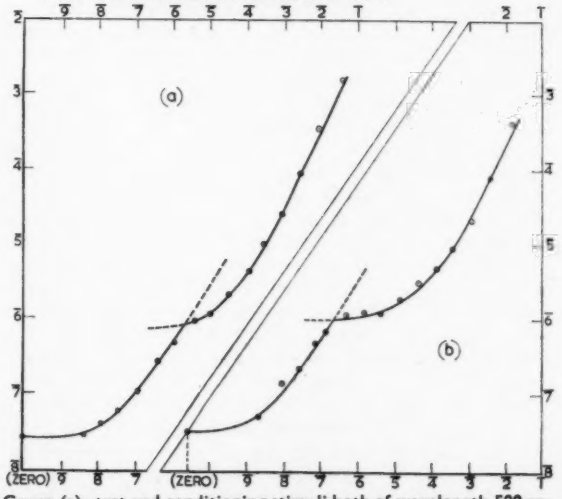
the three cone mechanisms. The rod sensitivity as normally determined cannot be so represented, but Hunt invokes a theory put forward by Dartnall (1948) that indicator yellow formed by the bleaching of visual purple in the rods may act as a filter material shifting the spectral sensitivity curve of the rods to longer wavelengths where it would be representable in the required manner. However, there is little evidence in favour of this view.

In later work, now in course of publication, by Hunt, it has been shown that the increase in the saturation of colours with increase in the conditioning brightness takes place even when a 1° foveal test field is used which may be considered to lie within the rod-free area of the retina. Thus Hunt has established that some factor other than the intrusion of rod vision produces desaturation of colour at low brightness levels.

5. Separation of the Rod and Cone Mechanisms by the Two-Colour Threshold Technique

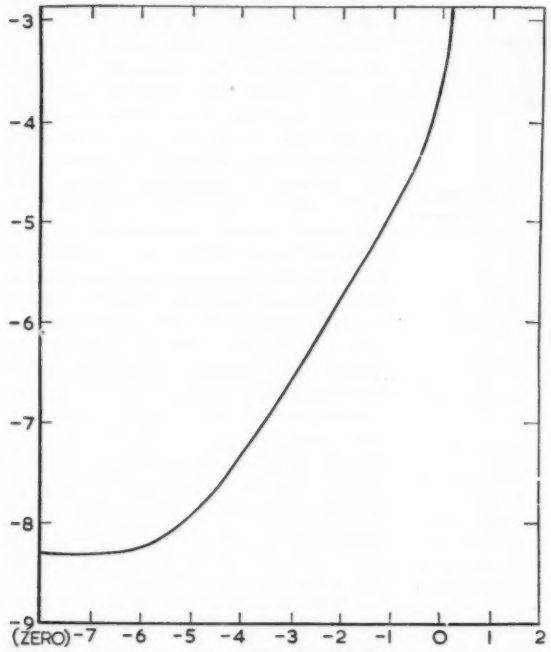
It has long been known that if certain visual capacities—visual acuity, contrast sensitivity, critical flicker frequency—are determined at various brightness levels (white light), a plot of the measured quantity against brightness gives a curve falling into low and high brightness sections attributable, respectively. to the rod and cone mechanisms. The possibility of separating visual mechanisms in this way is utilized in the two-colour threshold technique, with the important extension that, while the level is determined by a conditioning stimulus of one colour, the sensitivity is measured with a test stimulus generally of a different colour. The rod-cone separation is illustrated in fig. 3, of which curve (a) shows the increase in the increment threshold for a 1° yellow test stimulus viewed at 5° from the fovea on a yellow conditioning stimulus, both test and conditioning intensities being expressed in energy units and plotted on logarithmic scales. This curve may be interpreted on these lines. The lower branch continued in the way indicated by the broken line is assumed to represent the increase of increment threshold with conditioning intensity which would be obtained if the rod mechanism were acting alone, and similarly the extended upper branch corresponds to the cone mechanism acting alone. If the two mechanisms do not interfere with each other when both are acting, the test stimulus will be first seen at each conditioning intensity by means of the mechanism with the lower threshold and a curve corresponding to that observed will result. There may be some summation of rod and cone response when the thresholds are equal but the fairly sharp transitions shown by experimental curves suggest that such effects are small. If the rod and

FIG. 3. VARIATION OF THE INCREMENT THRESHOLD FOR 5°-PARAFOVEAL VISION, WITH INCREASE OF THE CONDITIONING INTENSITY



Curve (a): test and conditioning stimuli both of wavelength 580 mμ.
Curve (b): test stimulus of wavelength 580 mμ., conditioning stimulus of wavelength 500 mμ. (Stiles, 1939)
Abscissae: logarithm of the conditioning intensity in ergs/sec. deg.²
Ordinates: logarithm of the increment threshold expressed in ergs/sec. deg.² (test stimulus: 1° square, exposed for 0.063 sec.)

FIG. 4. VARIATION OF THE INCREMENT THRESHOLD WITH CONDITIONING INTENSITY FOR THE ROD MECHANISM ALONE: 5°-PARAFOVEAL VISION, 1°, 0.063-sec. TEST STIMULUS (Stiles, 1939)



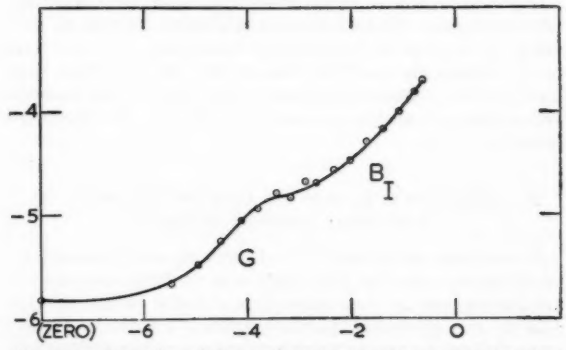
Abscissae: logarithm of the conditioning intensity in arbitrary units
Ordinates: logarithm of the increment threshold in arbitrary units

cone mechanisms have unique and stable relative spectral sensitivity curves, the effect of changing test and conditioning wavelengths on the curve relating increment threshold and conditioning intensity can be predicted. A change to a test wavelength for which the rod mechanism is more sensitive will lower the rod threshold by the same factor at all conditioning intensities, so that the rod branch will be moved bodily downwards in the diagram by an amount determined by the sensitivity change. A change of conditioning wavelength, on the other hand, will alter the effectiveness of the conditioning stimulus and this will displace the rod branch bodily to the left or right in the diagram by an amount depending on the rod sensitivity at the new and old conditioning wavelengths. A similar argument may be applied to the cone branch, the displacements being generally different, however, because of the different spectral sensitivities of rod and cone mechanisms. The effect of changing the conditioning wavelength from 580 mμ. to 500 mμ. (fig. 3, curves (a) and (b)) shows the increased lateral separation of the branches resulting from the greater sensitivity of the rods, and the rather smaller sensitivity of the cones to wavelength 500 mμ. Experiment has established that the rod branch maintains its shape and moves in the way required by the above picture, and in particular, that the same relative spectral sensitivity curve determines the effects of changing the test and conditioning wavelengths (Flamant & Stiles, 1948). The use of red conditioning stimuli enables the shape of the rod branch to be determined up to intensity levels which raise the threshold by a factor of some 100,000, and reveals a sharp upshoot of the curve at the highest intensities that may be caused by a kind of saturation of the rod mechanism whereby it is incapable of increased response, however much the stimulus be raised (fig. 4). Although the rod branch is invariant in shape for different test and conditioning colours, it varies with the size of the test stimulus and with the position of the latter in the extrafoveal retina (Cabello & Stiles, 1950). Such variations are inconsistent with the view, put forward particularly by Hecht, that the shape is determined solely by the progressive exhaustion of photosensitive material (visual purple) in the rods as the conditioning intensity is raised (see Hecht, 1938). It now appears probable that the extent to which the end-organs are grouped together by the neural system of the retina may be different at different intensity levels and at different locations, and may help to determine the shape of the rod branch. It has also been suggested that statistical fluctuations in the number of light quanta absorbed by the photosensitive substance, by setting a lower limit to the stimulus difference distinguishable from random changes—a limit down to which the rest of the mechanism is supposed to work—may be the main factor in determining the shape (De Vries, 1943).

6. Threshold Analysis of the Cone Mechanism

Returning to the effect of colour on the increment threshold, it might be expected that the cone branch would duplicate the behaviour of the rod branch, except that the relative spectral sensitivity curve would be a different one—displaced to longer wavelengths presumably. Experiment has shown, however, that the cone branch does not preserve its shape unchanged as test and conditioning wavelengths are altered, and cannot be regarded as a property of a single mechanism in the same way as the rod branch. The threshold properties of extrafoveal cone vision have not been much studied, partly because the presence of the very sensitive rod mechanism often masks the cone response. In most subjects the rod mechanism is absent from a foveal area of some 1° to 2°

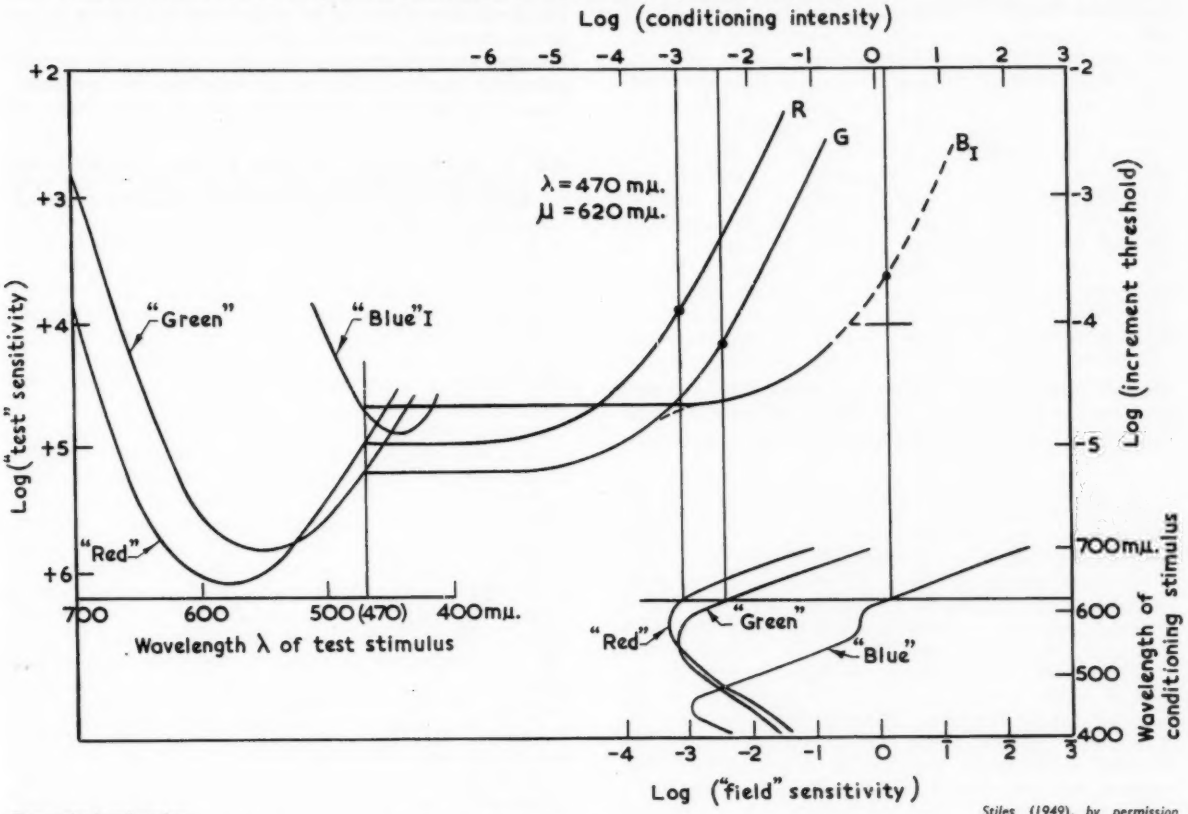
FIG. 5. VARIATION OF INCREMENT THRESHOLD FOR FOVEAL VISION, WITH INCREASE OF CONDITIONING INTENSITY: WAVELENGTHS OF TEST AND CONDITIONING STIMULI, 475 mμ. AND 550 mμ. RESPECTIVELY



Abscissae: logarithm of the conditioning intensity in ergs/sec. deg.^{2}
Ordinates: logarithm of the increment threshold in ergs/sec. deg.^{2}

G, B_I: branches of the curve attributed to different mechanisms

FIG. 6. DIAGRAM SHOWING THE WAY IN WHICH THE CURVE RELATING INCREMENT THRESHOLD AND CONDITIONING INTENSITY FOR DIFFERENT TEST AND CONDITIONING WAVELENGTHS (λ AND μ RESPECTIVELY) IS DETERMINED FROM THREE COMPONENT CURVES R, G, B_I BELONGING TO "RED", "GREEN" AND "BLUE" (B_I) MECHANISMS RESPECTIVELY (Stiles, 1939, 1949)



Stiles (1949), by permission of Documenta Ophthalmologica

Top right-hand scales:
Abscissae: logarithm of the conditioning intensity in ergs/sec. deg.^{2}
Ordinates: logarithm of the increment threshold in ergs/sec. deg.^{2}

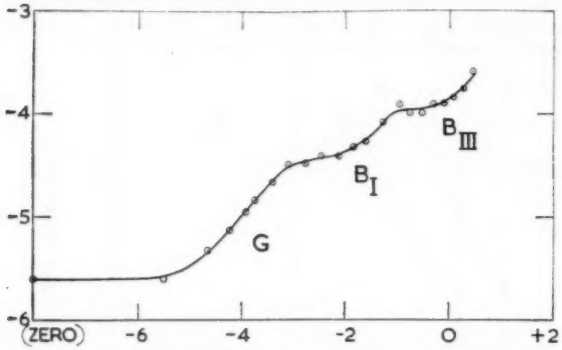
Bottom right-hand scales:
Abscissae: logarithm of the sensitivity of the mechanisms, expressed as the reciprocal of the intensity of the conditioning stimulus, in ergs/sec. deg.^{2}, necessary to raise the increment threshold of the appropriate mechanism by the factor 10 ("field" sensitivity)
Ordinates: wavelength of conditioning stimulus (mμ.)

Left-hand scales:
Abscissae: wavelength of the test stimulus (mμ.)
Ordinates: logarithm of the sensitivity of the mechanisms, expressed as the reciprocal of the increment threshold, in ergs/sec. deg.^{2}, of the appropriate mechanism for a conditioning field of zero intensity ("test" sensitivity)

Note: the small horizontal side-branch of curve B_I represents the initial part of a further branch which has since been assigned to an additional mechanism—B_{III}.

diameter, and increment thresholds measured with a test stimulus confined to this area show all the characteristics of cone vision. The most interesting result obtained by applying the two-colour technique to foveal vision is the demonstration that for certain combinations of test and conditioning wavelengths the curve relating the increment threshold and the conditioning intensity falls into two and sometimes three branches, and that these branches are displaced when the test and conditioning wavelengths are changed as though they belonged to different mechanisms, each with its own spectral sensitivity curve. An example of the splitting of a foveal curve into two branches is shown in fig. 5. A crucial question is whether the lower branch could be a rod branch, but the perception of colour near the threshold, the absence of any long adaptation delay in reaching the threshold if the eye is given a strong preliminary light exposure, and the marked dependence of the threshold on the point of entry of the rays in the pupil (a typical cone property), all indicate otherwise. As test and conditioning wavelengths are changed the two branches shown in fig. 5 are displaced and from the displacements

FIG. 7. VARIATION OF INCREMENT THRESHOLD FOR FOVEAL VISION, WITH INCREASE OF CONDITIONING INTENSITY: WAVELENGTHS OF TEST AND CONDITIONING STIMULI, 476 mμ. AND 578 mμ. RESPECTIVELY



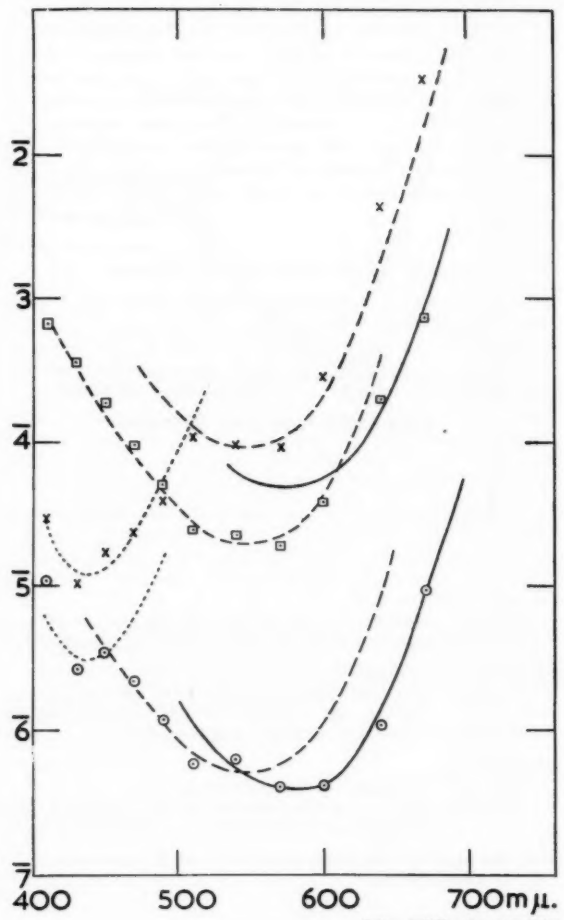
Abscissae: logarithm of the conditioning intensity in ergs/sec. deg.^{2}
Ordinates: logarithm of the increment threshold in ergs/sec. deg.^{2}
G, B_I, B_{III}: branches of the curve attributed to different mechanisms

it can be deduced that the upper and lower branches are attributable respectively to mechanisms with maximum sensitivities in the blue (440 mμ. approx.) and in the green (540 mμ. approx.), (B_I and G mechanisms). For test stimuli of wavelength greater than about 520 mμ., the B_I branch has moved completely above the G branch and is no longer in evidence. But the experimental curves for these longer test stimulus wavelengths indicate that another mechanism, with maximum sensitivity at about 580 mμ., must be present although a clear-cut separation like that between the B_I and G branches is not obtained. The manner in which the curve of increment threshold against intensity for any test and conditioning wavelengths is supposed to be derived from three branches belonging to B_I, G, and R mechanisms, respectively, is illustrated in fig. 6. When this diagram was put forward (Stiles, 1939), it was already known that the B_I branch showed an anomalous flattening out at high intensities of orange or red conditioning stimuli. Later measurements at still higher intensities make it probable that this flattening is the initial part of another branch attributable to a mechanism B_{III} which like B_I has maximum sensitivity at about 440 mμ., but a considerably lower sensitivity in the green and red (fig. 7). The diagram of fig. 6 was based mainly on measurements for one subject, and measurements for others have since shown that there is still another branch (B_{II}) corresponding to a mechanism with maximum sensitivity in the blue but with much higher sensitivity than B_I in the green and red. This mechanism is also probably present for the subject of fig. 6, but with a lower absolute sensitivity so that it is almost completely masked by B_I. The spectral sensitivities of B_{II} and B_{III} have not yet been completely determined. The investigation of the mechanisms G and R by the threshold method is difficult because of the extent to which their spectral sensitivity curves overlap, but here too the latest work suggests that there may be a secondary G and a secondary R mechanism coming in, in each case at high intensities of conditioning stimuli having wavelengths well removed from the wavelength of maximum sensitivity of the particular mechanism (Stiles, unpublished results).

7. Relation of the Threshold and Colour-Matching Mechanisms

What is the status of the mechanisms postulated in the above discussion? Their existence is merely inferred from the appearance of more or less well-defined branches in curves of the increment threshold against conditioning intensity—branches which preserve a certain individuality as the stimulus colours are changed. The end-organs of these mechanisms—presumably all cones—must be the same as those concerned

FIG. 8. VARIATION OF THE FOVEAL INCREMENT THRESHOLD WITH THE WAVELENGTH OF THE TEST STIMULUS UNDER THREE CONDITIONINGS



Stiles (1949), by permission of Documenta Ophthalmologica

Abscissae: wavelength of the test stimulus (mμ.)
Ordinates: logarithm of the increment threshold in ergs/sec. deg.^{2}

Conditioning stimulus:
- ⊙ zero field
- x high intensity red field
- ⊡ 2 sec. after extinction of red field

——— "red" mechanism
— — — — "green" mechanism
- - - - - - - - "blue" mechanism

in colour-matching, and their spectral sensitivities must be composites of the end-organ sensitivities although not necessarily linear combinations of these. In fact, however, the spectral sensitivities of the three main mechanisms—B_I, G, R—can be represented fairly well by linear combinations of the experimental curves x_λ, y_λ, z_λ of monocular colour-matching. Furthermore, they are placed along the spectrum in much the same positions as the fundamental spectral sensitivity curves derived by Wright (1934) and Walters (1942) by the binocular matching method. Insufficient is known of the spectral sensitivities of other threshold mechanisms to say whether they are also linear combinations of x_λ, y_λ, z_λ. It should be noted, however, that B_{III} has been observed only with very intense conditioning stimuli, after exposure to which monocular colour-matches break down. Superficially the suggestion that more than three mechanisms are separable by threshold methods may seem at variance with the statement that in binocular colour-matching only three independent mechanisms are involved. Apart from the loop-hole in this last assertion already mentioned in the first paragraph of section 4, the qualification " independent " removes any actual conflict. To enlarge a little on the point, suppose a match of two stimuli is arrived at by adjusting them until the subject answers no to the three questions: do they differ in brightness, in hue, in saturation? Once this condition is reached, he will also answer no to such questions as: do they differ in yellowness, blueness, etc.? These questions are not mathematically independent of the first three and they are redundant for colour-matching but it is conceivable that the eye may contain mechanisms for answering such " dependent " questions about yellowness and so on, and that these mechanisms may be picked up by suitable methods.

In discussing the threshold mechanisms, no reference has been made to linkage of mechanisms which the binocular matching results appear to demand. This may be because the behaviour of the threshold mechanisms has not yet been sufficiently explored. But there is another possibility which springs from the fact that in one set of measurements the eye is in equilibrium, whilst in the other it is in a state of change. It appears that interaction effects are most marked in the latter condition. The results of an experiment which illustrates the dramatic change in threshold properties that can occur after the removal of a high intensity conditioning stimulation are shown in fig. 8 (Stiles, 1949). The three curves of increment threshold against test stimulus wavelength can be roughly analyzed into parts for which the B_I, G, and R mechanisms are responsible. It is seen that the high intensity red conditioning raises the threshold for B_I by a small amount, that for G by a much greater amount. The threshold for R is raised so much that no part of the observed curve, even in the red end of the spectrum, is assignable to R. But two seconds after the removal of the conditioning stimulus the G and R mechanisms have begun to recover in the expected manner—their thresholds diminish—while the threshold of the B_I mechanism has increased so much that it is higher than the G threshold down to the blue end of the spectrum, and no part of the experimental curve is assignable to B_I. This experiment may be compared with Wright's conclusion that yellow conditioning produces an abnormal depression of the sensitivity of the " blue " process as determined by the binocular method. It may well be that under conditions of steady adaptation the visual process presents simpler—and perhaps misleadingly simple—properties.

A form of the two-colour threshold technique has also been used by Willmer (1949, 1950) in studying the properties of the small central foveal area (some 30′ in diameter) which, in colour-normal subjects, exhibits the colour-matching characteristics of dichromatic—actually tritanopic—vision. For this purpose very small test stimuli (10′ in diameter) must be used and fixation is more critical than with the 1° test stimuli considered up to now. On the other hand, dichromatic vision should be simpler than trichromatic vision. The results obtained by Willmer seem to show that the conditioning effect of violet light is about the same for test stimuli of all colours but that red conditioning has a progressively greater effect as the wavelength of the test stimulus is reduced. The measurements are not yet sufficiently advanced to provide a scheme comparable with that of fig. 6, which is based on 1° test stimuli.

ACKNOWLEDGEMENTS

I am grateful to Professor W. D. Wright and to Mr R. W. G. Hunt for permission to reproduce figs. 1 and 2 respectively.

REFERENCES

Cabello, J. & Stiles, W. S. (1950) *An. Fis. Quim.* **46,** 251
Dartnall, H. J. A. (1948) *Brit. J. Ophthal.* **42,** 793
De Vries, H. (1943) *Physica,* **10,** 553
Flamant, F. & Stiles, W. S. (1948) *J. Physiol.* **107,** 187
Hecht, S. (1938) *La base chimique et structurale de la vision.* Hermann, Paris
Hunt, R. W. G. (1949) *Proc. phys. Soc. Lond.* B, **62,** 203
Hunt, R. W. G. (1950) *J. opt. Soc. Amer.* **40,** 362
Hunt, R. W. G. (1952) *J. opt. Soc. Amer.* **42,** 190
Polyak, S. L. (1941) *The retina: the anatomy and the histology of the retina in man, ape, and monkey, including the consideration of the visual functions, the history of physiological optics, and the histological laboratory technique.* University Press, Chicago, Ill.
Roaf, H. E. (1932) *Proc. roy. Soc.* B, **110,** 448
Stiles, W. S. (1939) *Proc. roy. Soc.* B, **127,** 64
Stiles, W. S. (1949) *Docum. ophthal.* **3,** 138
Stiles, W. S. & Crawford, B. H. (1933) *Proc. roy. Soc.* B, **113,** 496
Walters, H. V. (1942) *Proc. roy. Soc.* B, **131,** 27
Willmer, E. N. (1949) *J. Physiol.* **110,** 422
Willmer, E. N. (1950) *J. Physiol.* **111,** 69
Wright, W. D. (1934) *Proc. roy. Soc.* B, **115,** 49
Wright, W. D. (1936) *J. Physiol.* **87,** 23
Wright, W. D. (1937) *Proc. roy. Soc.* B, **122,** 220
Wright, W. D. (1946) *Researches on normal and defective colour vision.* Kimpton, London

STIMULATION OF THE RETINA WITH LIGHT FIELDS OF SMALL SIZE

L. C. THOMSON Ph.D.

Medical Research Council
Group for Research in the Physiology of Vision
Institute of Ophthalmology, London

1 Foveal tritanopia
2 Relationship of these findings to visual theories
3 Measurement of the spectral sensitivity of the foveal centre with a small field
4 Intensity discrimination
5 Work with very small fields
6 Conclusion
References

In the design of experiments which might reveal the way in which the visual system works, investigations, in the first three decades of the century, and in so far as the colour-perceiving mechanisms are concerned, were almost entirely limited to subjective procedures. An observer was asked to view the light delivered by some suitable optical system and to perform certain operations, such as the matching in colour and luminosity of two adjacent patches of light. This he did by moving suitable controls on the instrument. The visual sensations aroused by the light were used merely as indicators of the equality of stimulation of the eye and the measurements consisted of records of the amounts of the light stimuli necessary to produce this equality.

One disadvantage from which such subjective procedures suffer is that there are so many different ways in which the experiments can be carried out and this in turn leads to difficulty in comparing one set of data with another. Thus the tendency in design was to develop a set of standard conditions under which accurate measurements were possible. This has led to a system of precise specification for all colours as seen by the average eye, which has proved invaluable to industry.

One of the standard conditions used in colour experiments is to make the test light patch of such a size that it subtends 2° at the eye. This field might be divided in half to form two adjacent areas for a matching experiment, but its over-all size would not exceed 2° because this was considered to be the largest field of which the image would lie almost wholly on the rod-free area of the fovea when the eye looked directly at it. This field, if of adequate luminance, could then be used to investigate the functions of the cones and colour-perceiving mechanisms without interference from the rod receptors. The use of such a field for the measurement of colour, and in the control tests during the manufacture of coloured articles, has now become general.

From the theoretical point of view, however, little further information is now to be gained by adhering to the standard observing conditions and, although much has been learnt about the colour mechanisms when a 2° field is used, a great deal of new information has been obtained by restricting the field diameter to an angle of only 15′ of arc, i.e., 0.25°. Such restriction has the advantage that many fewer photoreceptors are stimulated by an image of this size, so that an over-all or bulk response which is called into play by the 2° field might show separation into its component parts. The information gained in this way might solve the many theoretical problems of colour vision and in this article an attempt is made to assess this contribution to visual theory.

1. Foveal Tritanopia

König (1894) was the first to notice that the moon's disk could be made to disappear if it was viewed through a blue filter with the centre of the central fovea. This experiment showed that the central region was much less sensitive than the rest of the fovea to blue light. In addition, coloured stimuli of small size and of a blue or green colour were easily confused if viewed with the centre of the fovea. Such blue-green confusions are found in a form of defective colour vision called tritanopia and so the name " foveal tritanopia " has been given to this defect.

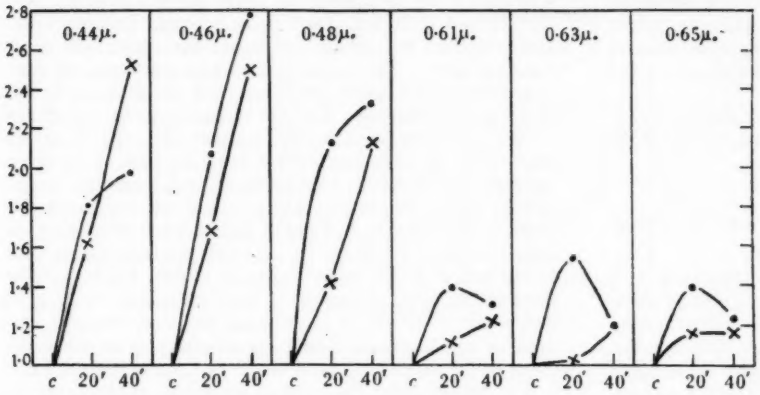
König's observations lay unheeded until 1944 when Willmer, who had developed a new theory of colour vision, rediscovered the blue-green confusions. In the following year they were investigated by Willmer & Wright (1945) on the Wright colorimeter. Figure 1 shows hue discrimination curves for the fovea which measure its ability to distinguish between adjacent wavelengths of the spectrum (Thomson & Wright, 1947). The abscissae record the wavelength at which the hue discrimination is measured and the ordinates the difference of wavelength, $\Delta\lambda$, between two halves of a 15′ field which enabled the observer just to detect a difference of colour. Since small values of $\Delta\lambda$ indicate relatively good hue discrimination, one can see from the lower of the three curves that, at the very centre of the fovea, wavelengths in the orange region of the spectrum (600 mμ.) can be readily distinguished

FIG. 1. THE HUE DISCRIMINATION OF AN OBSERVER, MEASURED WITH A 15′ FIELD

Abscissae: wavelength (mμ.)
Ordinates: $\Delta\lambda$ (the difference of wavelength between two small fields, required at any wavelength for the observer to be able just to see a difference of colour) (mμ.)

Curves shown refer to (i) very centre of the fovea, (ii) patch displaced by 20′ from the centre and (iii) patch displaced by 40′ from the centre.

FIG. 2. INSENSITIVITY OF THE FOVEAL CENTRE TO BLUE LIGHT



Abscissae: retinal position of the stimulus patch (c, centre of the fovea; 20′ and 40′, displaced by 20′ and 40′ from the centre)

Ordinates: sensitivity of the retina (number of times by which the sensitivity of the displaced patch exceeds that of the centrally placed patch)

—×—×— } results for two observers
—●—●— }

from one another, but that in the blue and green region (460–540 mμ.) discrimination is poor. For other small areas of the fovea, displaced from the centre through angles of 20′ and 40′, the general shape of the curves is similar but there is an over-all deterioration of discrimination. These curves are similar to those found in persons with a defect of colour vision known as tritanopia (Wright, 1952).[1]

The insensitivity of the foveal centre to blue light first noticed by König (1894) was further investigated by Thomson & Wright (1947). The sensitivity, relative to the central position, of a small area of the fovea, 15′ in extent, was measured at places displaced through angles of 20′ and 40′ from the centre. The results are shown in fig. 2. Here each pair of curves refers to a single wavelength and the abscissae give the retinal position at which the measurements were made. The ordinates record the number of times by which the sensitivity of the displaced patch exceeds that of the centrally placed patch. The two curves refer to different observers. For blue wavelengths (440 mμ., 460 mμ. and 480 mμ.) considerable insensitivity of the centre of the fovea was found, whereas for red light (610 mμ., 630 mμ. and 650 mμ.) there were scarcely any significant changes.

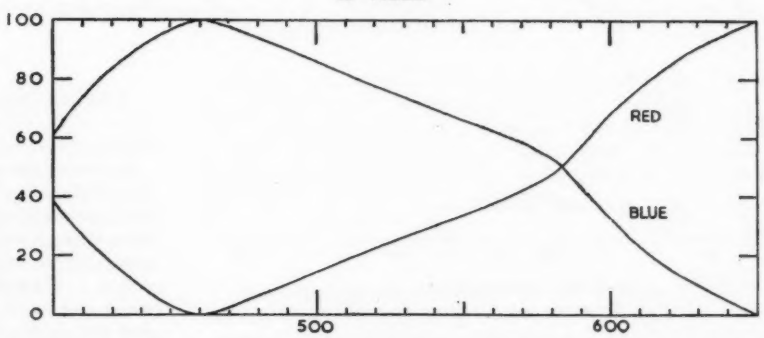
A further matching experiment carried out by Thomson & Wright (1947) gave the results shown in fig. 3 and confirmed the earlier observations of the same kind made by Willmer & Wright (1945). Here the ordinates give the proportions of a red and a blue light (wavelengths, 650 mμ. and 460 mμ.) which, when mixed together and viewed by the observer in one half of the 15′ field, would match in colour and luminosity each wavelength of the spectrum illuminating the adjacent half of the field. The abscissae give the wavelength of the monochromatic stimulus occupying this second half of the field. Only two matching stimuli were required to mimic all the colours of the spectrum and thus vision was said to be dichromatic. One interesting feature of the results is that the same proportions of red and blue light are required to match both wavelengths 515 mμ. and 425 mμ. For normal vision these wavelengths are coloured blue-green and violet respectively, and yet under the observing conditions of these experiments they can be matched completely merely by adjusting their relative intensities.

2. Relationship of these Findings to Visual Theories

Ever since Hermann von Helmholtz's mathematical statement of Thomas Young's hypothesis of colour perception, it has been known that the number of independent stimuli which must be mixed together in various amounts to mimic the sensations aroused by all other colours is equal to the number of independent mechanisms which exist somewhere in the visual pathway for the perception of colour. Since experiment has shown that the number of stimuli so required in the normal eye, when a 2° field is used, is three, vision is said to be trichromatic and it is known, therefore, that at some stage in the perceptual process there are only three independent mechanisms active. During the last century the validity of this remark has been repeatedly questioned, and for a full discussion of the case for and against it a review by Thomson (1952) should be consulted.

If, however, we accept the correctness of the above statement for the moment, a similar argument leads to the belief that there are only two colour-perceiving mechanisms in the central fovea, because under observing conditions which

FIG. 3. RESULTS OF A COLOUR-MATCHING EXPERIMENT WITH A 15′ FIELD



Abscissae: wavelength (mμ.)

Ordinates: the proportions of red and blue light (650 mμ. and 460 mμ. respectively) required to match each wavelength of the spectrum

The curves shown refer to the centre of the fovea.

[1] See also Wright, p. 36 of this Bulletin.

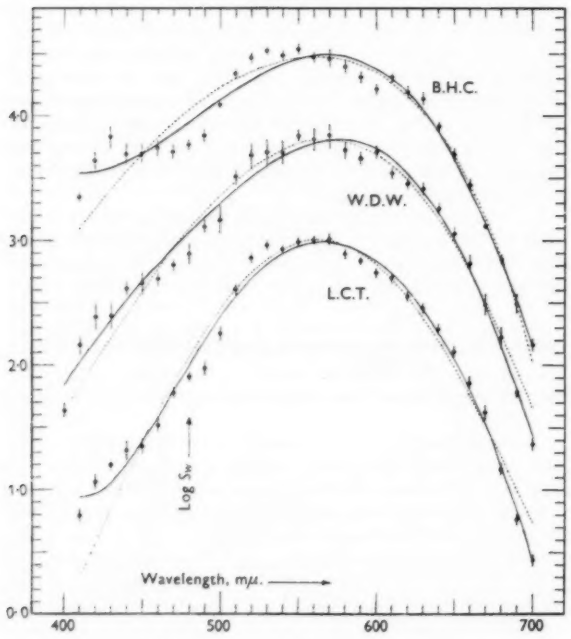
display foveal tritanopia only two independent stimuli are required to provide perfect matches for all colours. Since the retina is less sensitive to blue and violet light, one might suppose that the normal pathway for perceiving these colours is absent and that two others are left in a normal condition for appreciating red, yellow and green.

Thus the evidence so far leads to the postulate that the centre of the fovea in the normal eye has but two colour-perceiving mechanisms.

3. Measurement of the Spectral Sensitivity of the Foveal Centre with a Small Field

Stiles & Crawford (1933) were the first to obtain accurate spectral sensitivity data with a field of small size (40′ × 20′). In their experiments arrangements were made to light-adapt, by means of a large field, the area of the retina on which the image of a test stimulus of small size and known wavelength was imposed. The energy of the test flash required to make it just visible to the observer was recorded and the sensitivity taken as the reciprocal of this energy. Other measurements of a similar kind but in the dark-adapted eye have been made with a 15′ semicircular field by Thomson (1951). Figure 4, on the ordinates of which is given the logarithm of the sensitivity for various wavelengths of the spectrum, shows the

FIG. 4. THE SPECTRAL SENSITIVITY OF THE FOVEAL CENTRE



Abscissae: wavelength (mμ.)
Ordinates: log sensitivity (Log S_w)

The curves are mathematical functions of two kinds (the continuous and the dotted line curves respectively), fitted to the points and from which the points deviate significantly.
(The curves are artificially displaced one above the other.)
B.H.C., W.D.W., L.C.T.: observers

results obtained by Stiles and Crawford (B.H.C.) together with those for two observers with normal colour vision (L.C.T. and W.D.W.). Each point is shown together with a small vertical bar which represents the magnitude of its standard error. The curves are smooth mathematical functions fitted to the points and from which the data can be seen to deviate significantly. The sets of data have been displaced vertically relative to each other so that they may be clearly seen. Stiles & Crawford (1933) were the first to show the secondary maximum at wavelength 610 mμ., and other irregularities which are seen at 440 mμ. and 480 mμ. were first observed by Gibson & Tyndall (1923) when measuring the sensitivity with a 2° field. In addition there is a further peak at 520 mμ., first noticed by Thomson (1951). All these inflections in the spectral sensitivity data are either made more prominent or found by performing the measurements with fields of restricted size. Similar irregularities in the sensitivity curve of the fovea have been found by Crozier (1950).

It is possible that the over-all sensitivity of the fovea, as measured in the above experiments, is compounded from the individual sensitivities of a number of colour mechanisms. If so, then the irregularities in the total sensitivity curve could be evidence of its composite nature.

On the other hand, in all such sensitivity experiments the light energy from which the sensitivity of the eye was calculated has been measured at the cornea and in reality the light must pass through the optical media of the eye and finally through the inner layers of the retina itself before striking the receptors. In its passage it must suffer absorption which might be preferential for the wavelengths corresponding to the hollows between the raised portions of the curves in fig. 4. Adequate correction for this absorption might remove some of the irregularities.

Unfortunately, the data available for correction of the curves are poor and one cannot say certainly whether the irregularities are due to the appearance of the separate sensitivities of the colour mechanisms or not. However, lowering of the luminance at which the spectral sensitivity data were obtained increased the prominence of the secondary maximum at 610 mμ. (Thomson, 1947), so that in this case the irregularity of the over-all spectral sensitivity curve would seem to indicate imperfect fusion of two other sensitivities. If all the irregularities are due to this cause and the individual sensitivities are single peaked functions, then the mechanisms active in the central fovea would appear to be four or five.

4. Intensity Discrimination

A further experiment performed by Thomson (1949) on the ability of the eye to differentiate between various intensities of light of given wavelength also has a bearing on visual theory.

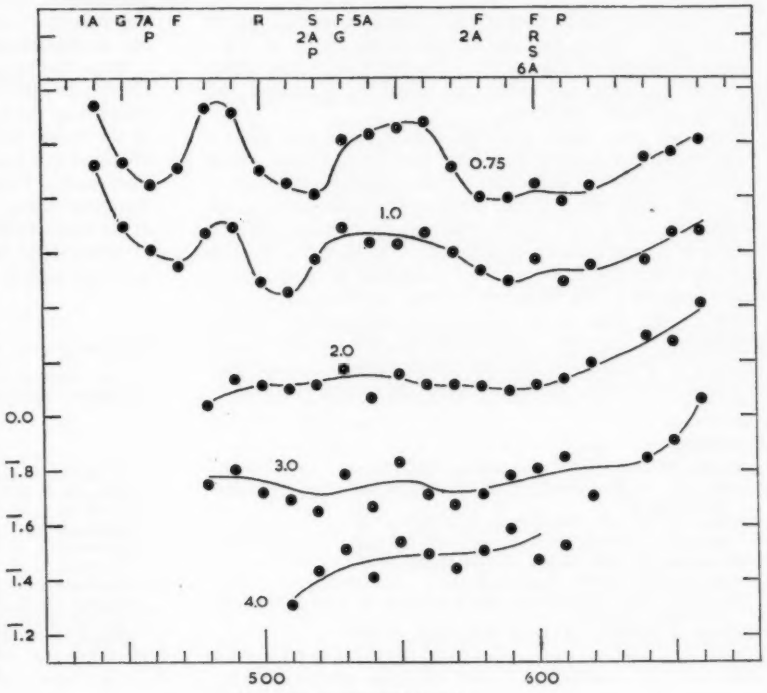
In these experiments the two halves of the 15′ field were both illuminated by light of similar wavelength. The lower half of the field acted as a comparison patch and was of various intensities, from that necessary to reach the threshold of vision to the maximum which the optical instrument would deliver. The observer could control the intensity of the upper patch and made his adjustments so that this field appeared just perceptibly brighter than the lower comparison patch. The difference of intensity ΔI was then expressed as a fraction of the intensity I at which the measurement was made. The logarithms of this quantity are plotted as ordinates in fig. 5, and small values indicate that the eye has good intensity

discrimination. Each full line curve passes through values of log ($\Delta I/I$) which appear of equal luminance to the eye. The upper curve is, at each wavelength, only very little above the threshold of vision and each curve below this was obtained at progressively greater values of luminance. The vertical separation of the curves is artificial and in the actual results all five curves overlap one another.

From fig. 5, log ($\Delta I/I$) seems to be independent of wavelength at high light levels (lower curves of the figure), but when the test patch has an intensity near the threshold of vision the eye is able to distinguish small differences of intensity more easily at some wavelengths than at others. These minima in the curve are at wavelengths of 460 mμ., 510 mμ. and 610 mμ. approximately.

The explanation of these minima rests upon the correlation of the findings with the results of some electrophysiological work by Granit (1947) on animals such as the rat, frog, etc. The action potentials from single fibres of the optic nerve were recorded and their sensitivities to light of various wavelengths measured. For certain preparations the sensitivity curves found were narrow in shape; Granit called these types of elements "modulators", and suggested that they were associated with the ability to discriminate between colours. The letters and numbers printed at the top of fig. 5 indicate the spectral positions of the maxima of these sensitivity curves, which were derived by a number of different methods. There is some agreement between these positions and the minima in the intensity discrimination data for the lowest light level. The ability to distinguish between small differences of intensity thus may depend at low light levels on receptors similar to the modulators found by Granit in animals. If so, then the curves of fig. 5 would indicate the activity of more than two mechanisms at the foveal centre.

FIG. 5. CURVES RELATING THE VALUE OF THE LOGARITHM OF THE RATIO $\Delta I/I$ TO THE WAVELENGTH FOR SEVERAL LEVELS OF BRIGHTNESS



Abscissae: wavelength (mμ.)
Ordinates: logarithm of the ratio ($\Delta I/I$)

The scale is correct for the highest brightness only, which is plotted at the foot of the diagram. The curves are labelled 0.75, 1.0, etc., in order of increasing brightness, but the numbers are not brightness values, merely labels. Curves labelled 1.0, 2.0 and 3.0 have each been displaced upwards by 0.4 log units and the curve labelled 0.75 by 0.2 log units. The letters at the top of the figure indicate the positions in the spectrum of maximum modulator activity as found by Granit (1947):

F: frog R: rat
G: guinea-pig S: snake

A, together with a cipher: maxima found in adaptation experiments. The ciphers (1, 2, 5, 6, 7) indicate the number of experiments yielding a maximum at each indicated spectral position.

P: maxima found by polarization

5. Work with Very Small Fields

A large number of experiments with light patches of very small size (a few minutes of arc only) have been undertaken by Hartridge (1947, 1950a, 1950b) and are described in another article in this issue.[2] Hartridge believes that the peculiarities of vision called here "foveal tritanopia" can be produced by viewing a sufficiently small field anywhere on the retina and also by lowering the light intensity in a field of larger size. His experiments suggest that only two mechanisms are operative when the eye is viewing a very small field: one of these has a maximum sensitivity at wavelength 500 mμ. and the other has two maxima of sensitivity, one at 420 mμ. and the other at 650 mμ. Thus the number of mechanisms postulated as active by Hartridge agrees with the number suggested by Willmer & Wright (1945). This conclusion is not supported by the work of Thomson (see above, section 3), which seems to require at least three—and possibly more—mechanisms to be present in the fovea and to be active under conditions revealing foveal tritanopia.

[2] See Hartridge, p. 31 of this Bulletin.

6. Conclusion

Discrepancies of interpretation, such as those revealed in the previous sections, have been common in the last hundred years and have led some to believe that there is an evanescent quality in the evidence presented by visual physiologists. One point of view, namely, that the mechanisms present must be limited and must correspond in number to the matching stimuli required to give unique matches in the colour mixture

experiments, has been argued vigorously in opposition to the view that many more mechanisms than this limited number are required to explain all the phenomena of colour vision. The Young-Helmholtz trichromatic theory is an example of the first set of views, and the polychromatic theory of Hartridge an example of the second. The battle of words, which at times has been very bitter, has lasted for a century and appears no nearer resolution.

Thomson has shown, however, that if those who hold a multimechanism hypothesis argue that each system exists separately and acts independently from its fellows throughout the whole extent of the visual pathway from the retina to the brain, they hold a view which cannot be harmonized with the results of colour mixture experiments. Such results lead to the conclusion that there is a limit in number to two independent mechanisms (in foveal tritanopia) or three (in normal vision). This limit is a mathematical consequence of the experiment and the only way in which all views can be brought together is to postulate that the visual pathway is composed of stages of activity. In one of these the independent mechanisms are limited. In other stages the mechanisms may be more than three in number and so help to explain other phenomena, but even in these stages the colour-matching experiment reveals that some of the extra mechanisms are but combinations of their fellows.

Stage theories of vision, such as that proposed by Müller (1930), do offer the best chance of solving the very difficult problem of the way in which colour messages are transmitted to the brain, but so far a successful attempt to fit the data obtained by a small-field technique into such a scheme has not been made. One can say with some certainty that the classical argument about the number of mechanisms used by the eye in the perception of colour is now sterile and that the results obtained from the small-field experiments do not sway the issue one way or the other.

ACKNOWLEDGEMENT

I am grateful to the Physiological Society for permission to reproduce, from the *Journal of Physiology*, figs. 2 and 4.

REFERENCES

Crozier, W. J. (1950) *J. gen. Physiol.* **34,** 87
Gibson, K. S. & Tyndall, E. P. T. (1923) *Sci. Pap. U.S. Bur. Stand.* **19,** 131
Granit, R. (1947) *Sensory mechanisms of the retina: with an appendix on electroretinography.* Cumberlege, London
Hartridge, H. (1947) *Philos. Trans.* B, **232,** 519
Hartridge, H. (1950a) *Experientia,* **6,** 1
Hartridge, H. (1950b) *Recent advances in the physiology of vision.* Churchill, London
König, A. (1894) *S. B. Akad. Wiss. Berl.*, p. 577
Müller, G. E. (1930) *Z. Psychol.* Ergbd 17 & 18
Stiles, W. S. & Crawford, B. H. (1933) *Proc. roy. Soc.* B, **113,** 496
Thomson, L. C. (1947) *J. Physiol.* **106,** 368
Thomson, L. C. (1949) *J. Physiol.* **108,** 78
Thomson, L. C. (1951) *J. Physiol.* **112,** 114
Thomson, L. C. (1952) *Ophthal. Lit.* **6,** 3
Thomson, L. C. & Wright, W. D. (1947) *J. Physiol.* **105,** 316
Willmer, E. N. (1944) *Nature, Lond.* **153,** 774
Willmer, E. N. & Wright, W. D. (1945) *Nature, Lond.* **156,** 119
Wright, W. D. (1952) *J. opt. Soc. Amer.* **42,** 509

* * *

COLOUR VISION IN THE PERIPHERAL RETINA

R. A. WEALE M.Sc.

Medical Research Council
Group for Research in the Physiology of Vision
Institute of Ophthalmology, London

1 Bearing of peripheral measurements on theories of colour vision
2 Special problems of research in the periphery
  a The peripheral pathways
  b Technique
3 Experimental findings
  a Colour-naming and wavelength discrimination
  b Spectral sensitivity
  c Macular pigmentation
4 An appraisal
References

1. Bearing of Peripheral Measurements on Theories of Colour Vision

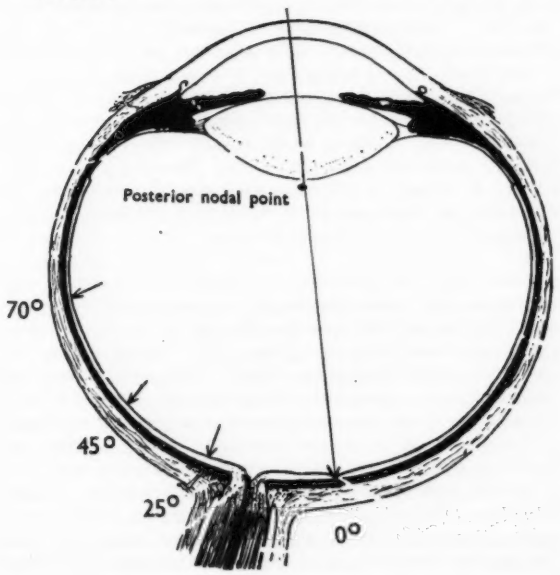
Schultze (1866) was the first to picture the human retina as two interdigitating systems: the daylight retina of the cones, and the twilight retina of the rods. Histological investigations (Østerberg, 1935 ; Polyak, 1941) have shown that the fixation area, located around the intersection of the visual axis with the retina, contains only cones. This rod-free area subtends an angle of about 50′ at the posterior nodal point (fig. 1). Outside this region, the number of rods increases, while that of cones declines. The fact that there exists a central scotoma in twilight vision supports the hypothesis of the functional division of the retina and has led to the duplicity theory, according to which rods mediate twilight vision, but cones daylight and (by implication) colour vision. Since twilight vision does not give rise to any marked sensation of colour, and the rod population predominates over that of the cones in the periphery, it has been inferred that colour vision is an exclusive property of the central fovea. This mistaken view is still widely accepted and even taught.

The mediation of colour sensations by peripheral regions, however, is to be expected since no part of the retina is cone-free. Colour sensations have been " demonstrated " in the periphery principally by means of the perimeter. The value of this instrument in clinical investigations is undoubted, but the numerical evaluation of a perimetric experiment is largely arbitrary and cannot be correlated with foveal experiments, in the majority of which a similarity or dissimilarity between two sensations is expressed in terms of stimuli. This, however, has not prevented attempts at correlation from being made. Indeed, the origin of the most serious rival of the Young-Helmholtz theory, Hering's hypothesis of opponent colours, can be traced to the observation of Hess (1889) that some spectral stimuli give rise to the same colour sensation at all retinal locations. Thus, whereas light of 600 mμ. appears orange foveally and yellow peripherally, light of 574.5 mμ. is invariably yellow. Similarly, spectral stimuli of $\lambda = 495$ mμ. (green) and $\lambda = 471$ mμ. (blue) produce spatially invariant sensations.

Testing the tristimulus theory (or trichromatic theory as it is usually called) with the results obtained by Hess is difficult. In accordance with the more modern formulations of this theory, the chromatic visual system possesses a triadic bottle-neck at some stage between the retinal receptor layer and the seat of perception. This theory has been postulated as a result of colour-matching experiments, carried out in the fovea. Similar experiments have not been carried out in the periphery. Conversely, it is hardly possible to carry out the " invariant sensation " experiment in such a tiny area as the fovea. It is therefore illogical to repeal the tristimulus theory because it does not predict an effect for the peripheral retina before it has been shown to be valid in this region, just as it is illogical to reject the Hering theory on the grounds of foveal experiments.

A more general argument against the tristimulus theory is derived from a study of the perimetric fields. Abney (1913) found that a mixture of monochromatic red and green, which matches a monochromatic yellow when seen foveally, has a smaller perimeter than the yellow. Hartridge (1949) thought that this observation invalidates one of the conclusions of the tristimulus theory, that is, that yellow is due to the equal stimulation of the long and medium wavelength mechanisms respectively. This criticism of the theory implies that the colour-mediating mechanisms of the periphery are exact—if diluted—replicas of those in the fovea. But such is not the case. The difference between fovea and periphery does not consist only in the density of the cone population: the post-receptoral structures differ and there are rods in the periphery. The larger perimeter for yellow can be explained easily if a desaturating effect of the rod mechanism is postulated

FIG. 1. A HORIZONTAL MERIDIONAL SECTION THROUGH THE RIGHT EYE-BALL, INDICATING THE POSITION OF THE POSTERIOR NODAL POINT AND PERIMETRIC ANGLES OF 0°, 25°, 45° AND 70°



(Lythgoe, 1931). But there is an even simpler explanation. In the first instance, the foveal match in Abney's experiment cannot have been perfect since the monochromatic yellow would have to be slightly desaturated to match the mixture. Let it be assumed, however, that a perfectly matched mixture still gives rise to the same result, i.e., it turns achromatic before the monochromatic light when the perimetric angle is increased. Owing to the enhanced sensitivity to light of short wavelengths in the periphery, the green component of the mixture will evoke a greater response from the " blue " mechanism, postulated by the tristimulus theory, than will the yellow or the red stimuli, the result being a desaturation of the mixture in conformity with the tristimulus theory. The present explanation is simpler than Lythgoe's because only the experimental datum of improved sensitivity to short wavelengths in the periphery is used.

These examples show the relation of the retinal periphery to questions of theory. Most of the peripheral work done in the past is irrelevant to the confirmation or contradiction of theories of colour vision. When a particular theory is being tested by investigations in the periphery, it is a prerequisite that the methods used for the fovea should, as far as possible, be preserved. The reason why this has not been done in the past is largely the difference in character between foveal and peripheral vision, particularly as regards form vision.

2. Special Problems of Research in the Periphery

a. The Peripheral Pathways

The foveal and extrafoveal regions differ most strikingly in visual acuity. If one letter in this line is fixated, hardly more than two or three words on either side of it can be distinguished. The immediate reasons for this are: (i) that the better image is formed in the fovea, and (ii) that the cone paths in the foveal pit are little interconnected and so make possible the provision of a mechanism for the appreciation of fine detail. Histological investigations (Polyak, 1941) have shown that the human periphery resembles the feline retina in its far-flung arborizations of receptor processes and connexions between rod and cone pathways. According to Polyak, the absence of midget bipolars in the periphery and their presence in the fovea may indicate their role in colour perception. Since colour can be perceived peripherally, however, Polyak believes that the peripheral cones possess the colour-mediating function which he denies to the foveal ones on account of their egg-like similarity to one another. This idea is interesting because it recognizes that more than one of the eight retinal structures can partake in the wavelength analysis which is required for the appreciation of colour.

Before we leave this discussion of the retina, the macular pigment should be mentioned. Schultze (1866) saw it only in primate eyes; since this suggests its association with well-developed foveae, one wonders whether it can be found in the eyes of birds like the golden eagle, whose foveae are very deep indeed (Pumphrey, 1948). Gullstrand (1906) has insisted that the appearance of the macular pigment in man is due to a post-mortem transudation of serum in the thinnest part of the retina, namely the " macular " region. Difficulties for this hypothesis are: (i) that the transudation would have to be well-nigh instantaneous, since Chevallerea and Polack (cf. Holm, 1922) observed a yellow coloration in the foveal region only eight minutes after enucleation of the eye, and (ii) that the absorption characteristics of serum differ from those of macular pigment extracts as obtained by Hanström (1940) for the macaque and by Wald (1949) for man. It is also difficult to understand why such a transudation should take place at all: the blood pressure drops long before the relative impermeability of the containing vessels. In general, the evidence for the belief that the macular pigment, found in dead eyes, is solely due to post-mortem changes is unconvincing. Recently, moreover, Brindley & Willmer (1952) reported measurements made on the living eye, which suggest that the peripheral retinal media absorb light of short wavelengths (blue and green) to a smaller extent than the foveal media. The absorption of light by the macular pigment is thought to affect the foveal sensitivity; since light passes the retina twice when the fundus is examined ophthalmoscopically, some difference in colour would be expected as between the fovea and periphery. Thus the uniform colour of the fundus, presumed in this work, and the importance of the macular pigment, are mutually inconsistent. Unfortunately, the measurements of Brindley and Willmer cease at 460 mμ., below which is the most interesting spectral region.

The next portion of the visual pathways which has been investigated in connexion with colour vision is the lateral geniculate body. Le Gros Clark (1949) has suggested that its six striations may be associated with fibres conducting impulses which produce red, green and blue sensations: striations 1, 4 and 6 belong to the contralateral, 2, 3 and 5 to the ipsilateral eye; 1 and 2 comprise " green", 3 and 4 " red", and 5 and 6 " blue " fibres. Having plotted the topography of a section of the lateral geniculate body in terms of retinal positions, Le Gros Clark found that it was at its thickest in the " foveal " region, tapering off towards the " periphery". The first four layers fused into two, a fact which Le Gros Clark associated with the dichromacy of the peripheral retina. Striations 5 and 6 diminished at a much slower rate than the others: this fact is correlated with the relatively good " peripheral " appreciation of blue. The scheme is attractive but Chacko (1949) believes that there may be four striations for each of the two primate eyes. Hartridge (1950a) has given some cogent criticisms of Le Gros Clark's interpretation of these anatomical data: in this simple form it cannot account for the fact that the perimetric fields for different colours are a function both of the intensity and of the areal subtense of the test stimulus. Moreover, no part of the lateral geniculate body can be associated with the transmission of rod impulses.

Lastly, one of the paradoxes of the structure of the visual pathways is the dissimilarity between the retinal projections of the fovea and periphery, respectively, on the visual cortex: while the fovea is so much smaller than the retinal periphery, the reverse applies to their projections (Holmes, 1944). This fact underlines the importance which we ascribe to messages received via the fovea as compared with the inferior value of peripheral impressions.

It is in this discrepancy that the difficulties of visual research on the periphery find their culmination. Of some 45 papers on colour vision published between 1920 and 1952 in Great Britain only 7 deal with the periphery. Every single one of the above factors militates against the investigation of the periphery: the aberrations in the dioptrics of the ocular media conspire against the use of small test fields, which is desirable for reasons given elsewhere.[1] The neural interconnexions in the periphery reduce the visual acuity and lead to a severe desaturation of colour. The presence of macular pigment

[1] See Thomson ,p. 50 of this Bulletin.

in the fovea aggravates the difficulty of comparing peripheral with foveal data. Last, but not least, the disparity between the foveal and peripheral projections on the visual cortex impedes fixation simultaneous with the observation of events taking place in the periphery.

b. Technique

Until Guild, Wright and Stiles had demonstrated the necessity of properly controlling experimental conditions in foveal measurements, the quality of most investigations left much to be desired. In peripheral work, in particular, some of the best-known earlier investigations (Hess, 1889) were carried out with pieces of coloured paper. This habit has persisted even to the present day; Bäuml & Engelhardt (1951) viewed Ostwald papers in their perimetric colour-naming tests, while Hartridge (1947) used Munsell papers in an investigation of the peripheral colour sensations as a function of luminance level and areal subtense. But spectral stimuli should be employed in peripheral as well as in foveal work because the experimenter benefits by having a greater range of potential stimuli and is enabled to give an adequate description of the stimulus. In his later work on the determination of the invariant colours, Hess (1889) used spectral stimuli. Abney's measurement of the photopic spectral sensitivity curve at a perimetric angle of 10° was carried out with his spectral colour-patch apparatus (Abney, 1913). The earlier work on the peripheral photopic sensitivity (cf. Parsons, 1924) was also carried out with spectral stimuli but suffers from the defect that the precise retinal locations were not stated and no definite fixation appears to have been maintained.

One of the most important factors has received but scant attention in the past, that is, the level of adaptation. Most of the work on the photopic peripheral sensitivity was carried out with stimuli whose luminance level was admittedly photopic by foveal standards, but the surrounding field was dark and thus the eye was not generally light-adapted. This fact has led to some results which are so remarkable that they will be given separate attention below. Thus the principal points of technique which have to be considered in sensory work can be summarized as follows: (i) adequate control of the stimulus; this includes the intensity and the spectral composition of the stimulus and the pupil size, and the points of entry in the pupil of the test beams; (ii) control of the localization of the stimulus, i.e., reliable fixation so that the perimetric angle can be measured; (iii) controlled light-adaptation.

A new and entirely different step in the technique of peripheral investigation has been taken by Motokawa (1949), who measured the threshold voltage required to produce a phosphene after previous illumination of a given retinal area by spectral lights. The inadequacy of the statistical evaluation of the data makes an assessment of the peripheral results difficult. This method may yield information about such physiological entities as latent periods of foveal and peripheral colour-mediating mechanisms.

3. Experimental Findings

a. Colour-Naming and Wavelength Discrimination

The modern descriptions of peripheral sensations are characterized by the awareness of the importance of the level of adaptation. The colorimetric measurements of Hunt (1950) have shown the effect of adaptation on foveal colour assessment. Though no comparable measurements have as yet been obtained for the periphery, the work of Wentworth (1930), Lythgoe (1931) and Weale (1951a) shows that light-adaptation has an even profounder effect on the periphery than on the fovea. Wentworth investigated the perimetric fields for white and four spectral stimuli (red, 672.5 mμ.; yellow, 581.5 mμ.; green, 522 mμ.; and blue, 468 mμ.). She found that remarkable interlacing took place when stimuli of equal energy were used. Abney (1913) had previously commented on the concentricity of the fields as determined with stimuli of equal luminance. Wentworth showed further that the fields were very nearly coextensive when their maximal extents were determined with a 1° test stimulus: in this case, of course, no energy limitations were imposed. The chromatic sensitivity, i.e., the reciprocal of the energy at which a colour was recognized as such, depended very greatly on the state of adaptation. In dark-adaptation the colour sensitivity decreased more rapidly with the perimetric angle (in the horizontal meridian). Moreover, there was an interesting change in the size of the fields: the maximum chromatic sensitivity tended to shift from the shorter to the longer wavelengths as the level of adaptation was reduced. This association between increased level of light-adaptation and improved visual functions at short wavelengths has been noted by several workers (cf. sections 3b and 3c).

Lythgoe carried out a colour-naming test for the light- and dark-adapted eye. Perimetric angles of up to 70° were investigated and light-adaptation was found greatly to enhance the saturation of colours viewed peripherally. Blue was the only colour which was recognized as such in both states of adaptation even in the far periphery, a finding recently confirmed by Bäuml & Engelhardt (1951). These qualitative observations are supported by measurements of the wavelength discrimination which show not only that this visual function depends on the luminance level particularly in the central (green) region of the spectrum, but also that it is at its best in the blue (Weale, 1951a, 1953). This latter point is illustrated in fig. 2b in which the just noticeable wavelength step $\Delta\lambda$ (ordinate) is plotted against the shorter wavelength of the step (abscissa) for the fovea and perimetric angles of 25° and 45° respectively. Excepting at short wavelengths (in the blue), the steps are much larger in the periphery than in the fovea, thus indicating a marked deterioration in wavelength discrimination. At 45° considerable red-green confusion is taking place (stimuli which appear green foveally may be orange when observed peripherally). Thus, the farther periphery tends to exhibit characteristics of dichromatic vision, a matter of great interest in view of the apparent tripartition of the sensitivity curves obtained for the same region (fig. 2a).

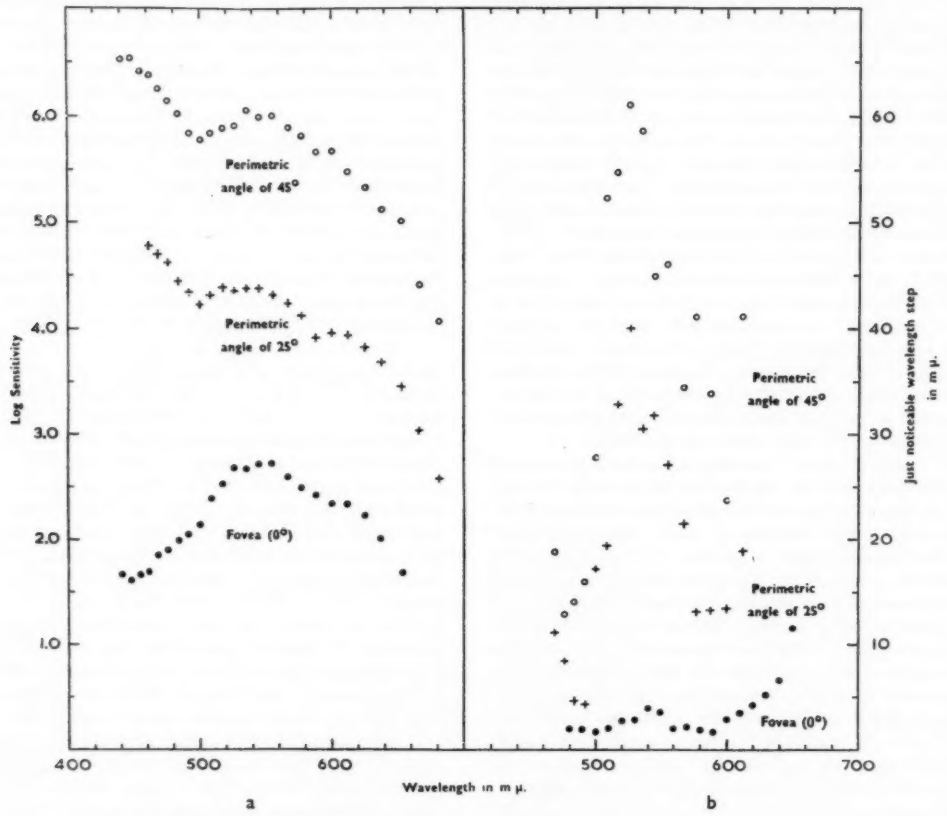
b. Spectral Sensitivity

The sensitivity is a measure of the ease with which a given visual sensation can be evoked. When it is measured for different wavelengths in terms of the reciprocal of the light energy producing such a sensation, a spectral sensitivity curve can be plotted. Whereas the differences between curves obtained for different levels of adaptation in the fovea are relatively small (Thomson, 1951), effective light- and dark-adaptation yield very different results in the periphery. The dark-adapted periphery gives rise to the scotopic curve when the stimulus level is low. If this level is high but a dim or zero surround is used, a curve is obtained which is intermediate between the scotopic and the foveal photopic curve (Tschermak,

1890; Wright, 1946). If, however, the level of adaptation of the surrounding field is also high and the test field subtends at the eye an angle of less than 1°, there obtains a remarkable state of affairs. The sensitivity curves obtained by Stiles & Crawford (1934), who investigated a perimetric angle of 5°, and by Weale (1951b, 1953), whose measurements

Moreover, the tripartition of the curve into divisions which cover the blue, green and orange parts of the spectrum respectively is curiously at odds with the data for wavelength discrimination, which suggest dichromatic vision in the periphery. This paradoxical result, paralleled by a similar state of affairs in the central fovea, awaits an explanation.

FIG. 2. SPECTRAL SENSITIVITY AND WAVELENGTH DISCRIMINATION AT DIFFERENT PERIMETRIC ANGLES



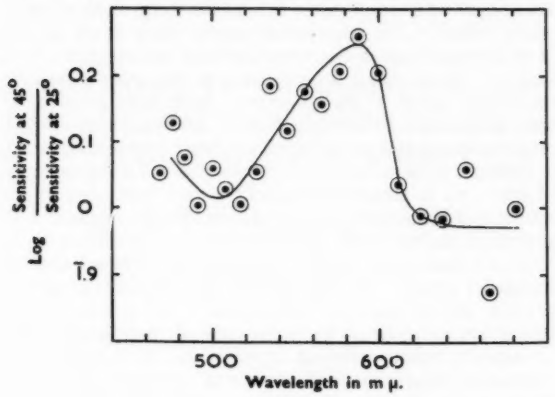
a: relative spectral sensitivity, measured at three perimetric angles
(The three sets of data are displaced vertically by arbitrary amounts to facilitate comparison.)

b: just noticeable wavelength step $\Delta\lambda$, measured at three perimetric angles

extended to 70°, exhibit a pronounced tripartition. This is illustrated in fig. 2a for the two perimetric angles of 25° and 45° and for a high luminance level. The logarithm of the sensitivity is plotted along the ordinate, the wavelength along the abscissa. The data for 25° and 45° are displaced vertically by an arbitrary amount to facilitate comparison. The foveal curve is also shown. It is at once apparent that the points follow similar trends at long wavelengths. However, there is a hump in the blue of the spectrum representing a marked enhancement of the sensitivity. This is so overwhelming that the maximum of the whole curve moves to this region.

Another point of interest is the improved absolute sensitivity to yellow light (about 570–590 mμ.), as the perimetric angle is increased. Using the phosphene method, briefly mentioned above, Motokawa (1949) claimed to have found a mechanism specific to the peripheral retina and dominant in the yellow part of the spectrum. Hartridge (1949) has suggested that such a mechanism was responsible for the photopic spectral sensitivity in the peripheral retina having its maximum at longer wavelengths than in the fovea, i.e., in the yellow. This view, however, is untenable since it is based on data to which the equal energy correction had not been

FIG. 3. IMPROVEMENT IN SENSITIVITY TO YELLOW LIGHT WITH INCREASE OF PERIMETRIC ANGLE



applied (cf. Parsons, 1924). It is none the less true that the sensitivity to yellow light does improve as the perimetric angle is increased. This is illustrated in fig. 3 where

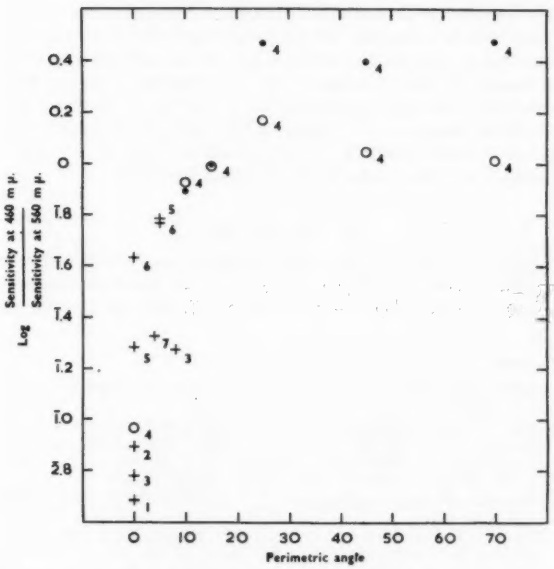
$$\log \frac{\text{sensitivity at } 45°}{\text{sensitivity at } 25°}$$

is plotted against the wavelength: the improvement shown in the region of 590 mμ. is absolute and significant when the data are tested statistically. The real cause of this improvement is still unknown.

c. Macular Pigmentation

Peripheral measurements have also been used to determine the extent (Wright, 1946) and amount of macular pigmentation. The question of whether or not any macular pigment is present in the living eye of man is one which is still debated (Dartnall & Thomson, 1949; Denton & Pirenne, 1950; Hartridge, 1950b; Tansley, 1950). To demonstrate its existence in man unequivocally has so far not been found to be possible. Wald (1949) has extracted a macular pigment from enucleated human eyes and identified it as a xanthophyll. Hanström (1940), on the other hand, obtained a negative Carr-Price reaction (for carotenoids) when using retinas from monkeys; his test for vitamin A also led to a negative result. In support of the belief that the macular pigment exists also in the living eye, Wald suggested that certain shape irregularities in foveal sensitivity curves were caused by the absorption bands of the macular pigment: a closer investigation of this comparison, however, shows that, while the bands partly account for the irregularities, they cannot do so completely (cf. Thomson, 1951). It is thought, furthermore, that a comparison of the photopic foveal and peripheral sensitivity curves may yield a quantitative estimate of the optical density of the macular pigment in situ. The basic assumption underlying this argument is that, under photopic conditions, the retinal sensitivity is independent of the region to which the stimulus is applied, or, more generally, as pointed out by Stiles (1949), that the spectral distribution of the relative sensitivity is constant. This implicit assumption was made by Wald (1949), who found an apparently quantitative agreement between the in situ density of the macular pigment and the value obtained from retinal extracts. If the assumption were valid, then any difference in the measured sensitivity could be ascribed to a differential filter effect due to pre-receptoral media. Figure 4 suggests that there is a marked difference in the actual sensitivity of the receptors and that this adds its effect to that of the reduction in the density of macular pigment as more peripheral retinal regions are tested. The graph represents a plot of the log ratio of the sensitivity at 460 mμ. to that at 560 mμ. (ordinate) against the perimetric angle (abscissa). The latter refers to a variety of angles in azimuth: according to Abney (1913), the sensitivity to white is independent of the azimuth. Wentworth (1930), however, observed marked differences. The present treatment necessarily ignores this fact. The choice of (log sensitivity$_{460\ m\mu}$ −log sensitivity$_{560\ m\mu.}$) as parameter is justified by the

FIG. 4. LOG $\frac{\text{SENSITIVITY AT 460 mμ.}}{\text{SENSITIVITY AT 560 mμ.}}$ PLOTTED AGAINST THE PERIMETRIC ANGLE



+ data obtained by observers numbers 1, 2, 3, 5, 6, 7

o data obtained by observer number 4, measured at a low luminance level

• data obtained by observer number 4, measured at a high luminance level

Observer number	Observer (initials)	Reference
1	L.C.T.	Thomson, 1949
2	W.D.W.	Thomson, 1949
3		Wald, 1949
4		Weale, 1953
5	B.H.C.	Stiles & Crawford, 1934
6	W.S.S.	Stiles & Crawford, 1934
7		Gilbert, 1950

following argument. The wavelength of 560 mμ. corresponds to the maximum of the photopic foveal sensitivity curve and 460 mμ. is the approximate wavelength of maximum change in the photopic sensitivity with perimetric angle (Weale, 1951b). Since a comparison of the data of numerous workers generally involves one arbitrary choice (for example, equating sensitivities at one wavelength), nothing is lost in choosing the above parameter: its advantage is that it tells at a glance how the sensitivity to a blue compares with that to a green stimulus. Thus, near the retinal centre the log values are negative, indicating that $\frac{\text{sensitivity (blue)}}{\text{sensitivity (green)}}$ is less than unity, i.e., the sensitivity to green is greater than that to blue light. The reverse applies to more peripheral regions. Now the interesting point in connexion with the question of macular pigment is that $\frac{\text{sensitivity (blue)}}{\text{sensitivity (green)}}$ increases at perimetric angles outside which the macular pigment is thought to exist. At large angles, moreover, there is a difference in this ratio according to whether the eye is adapted to a high or to only a moderate luminance level. It follows from these two points that a correct estimate of the in situ density of the macular pigment by comparing photopic peripheral with foveal curves is largely fortuitous because the distribution of the receptor mechanisms is not the same in the two cases. The difference due to luminance level is inexplicable on the basis of a macular filter and points to an enhanced activity of a "blue" mechanism under conditions of light-adaptation.

4. An Appraisal

Has the study of the retinal periphery contributed to the understanding of colour vision and, in particular, foveal colour vision? The answer is yes, although new problems have appeared. The state of adaptation, particularly indirect adaptation, plays a much larger role than in the fovea. Thus, light-adaptation raises the sensitivity to light of short waves (blue) to an unexpected extent: there is no known photochemical mechanism which explains this in terms of rod activity. Again, there is an increase in sensitivity specific to the yellow region of the spectrum. Both these effects have been predicted by Hartridge (1949), although their relative magnitudes are not as he predicted them. In general, there is sufficient evidence that colour vision in the accepted sense of the word is present throughout a very large part of the retina: it is possible to measure wavelength discrimination in peripheral regions in the same manner as in the fovea. The relative improvement of the sensitivity in restricted spectral regions produces in no way an argument against the general applicability of the tristimulus theory. As already stated, this theory is based on the sufficiency of three independent variables in colour-matching experiments. If the peripheral wavelength discrimination curves had ordinates lower than the foveal (fig. 2b), the possibility of four or more independent variables would have to be faced. But all the peripheral minimal wavelength steps are equal to or greater than the foveal ones and thus suggest quantitatively poorer but not qualitatively different colour vision in the periphery. The mere fact that the relative proportions of these variables, obtained foveally, may differ from those found for the periphery does not detract from the general validity of the above conclusion.

The problem of peripheral light-adaptation is not new but its quantitative description is only recent and offers promising lines of research. It seems to indicate that the eye is not just a pot full of pigments, but that its whole function can be elucidated only by a systematic inquiry into the nervous mechanisms thrown into activity by photochemical action.

References

Abney, W. (1913) *Researches in colour vision.* Longmans Green, London
Bäuml, J. & Engelhardt, A. (1951) *Z. Biol.* **104,** 307
Brindley, G. S. & Willmer, E. N. (1952) *J. Physiol.* **116,** 350
Chacko, L. W. (1949) *Arch. Ophthal., N.Y.* **42,** 402
Clark, W. E. Le Gros (1949) *Docum. ophthal.* **3,** 57
Dartnall, H. J. A. & Thomson, L. C. (1949) *Nature, Lond.* **164,** 876
Denton, E. J. & Pirenne, M. H. (1950) *Nature, Lond.* **165,** 304
Gilbert, M. (1950) *Proc. phys. Soc. Lond.* B, **63,** 83
Gullstrand, A. (1906) *v. Graefes Arch. Ophthal.* **62,** 1
Hanström, E. (1940) *Acta Ophthal., Kbh.* **18,** 21
Hartridge, H. (1947) *Philos. Trans.* B, **232,** 65
Hartridge, H. (1949) *Docum. ophthal.* **3,** 166
Hartridge, H. (1950a) *Recent advances in the physiology of vision.* Churchill, London
Hartridge, H. (1950b) *Nature, Lond.* **165,** 304
Hess, C. (1889) *v. Graefes Arch. Ophthal.* **35,** 1
Holm, E. (1922) *v. Graefes Arch. Ophthal.* **108,** 1
Holmes, G. (1944) *Proc. roy. Soc.* B, **132,** 348
Hunt, R. W. G. (1950) *J. opt. Soc. Amer.* **40,** 362
Lythgoe, R. J. (1931) *Brit. J. Ophthal.* **15,** 193
Motokawa, K. (1949) *J. Neurophysiol.* **12,** 467
Østerberg, G. (1935) *Acta Ophthal., Kbh.* Suppl. No. 6
Parsons, J. (1924) *Colour vision.* University Press, Cambridge
Polyak, S. (1941) *The retina: the anatomy and the histology of the retina in man, ape, and monkey, including the consideration of visual functions, the history of physiological optics, and the histological laboratory technique.* University Press, Chicago, Ill.
Pumphrey, R. J. (1948) *J. exp. Biol.* **25,** 299
Schultze, M. (1866) *Arch. mikr. Anat.* **2,** 165
Stiles, W. S. (1949) *Docum. ophthal.* **3,** 135
Stiles, W. S. & Crawford, B. H. (1934) *Proc. roy. Soc.* B, **113,** 496
Tansley, K. (1950) *Nature, Lond.* **165,** 524
Thomson, L. C. (1949) *Proc. phys. Soc. Lond.* B, **62,** 787
Thomson, L. C. (1951) *J. Physiol.* **112,** 114
Tschermak, K. (1890) *Pflüg. Arch. ges. Physiol.* **47,** 417
Wald, G. (1949) *Docum. ophthal.* **3,** 94
Weale, R. A. (1951a) *J. Physiol.* **113,** 115
Weale, R. A. (1951b) *J. Physiol.* **114,** 435
Weale, R. A. (1953) *J. Physiol.* **119,** 170
Wentworth, N. (1930) *Psychol. Monogr.* No. 40
Wright, W. D. (1946) *Researches in normal and defective colour vision.* Kimpton, London

THE ABSOLUTE SENSITIVITY OF THE EYE AND THE VARIATION OF VISUAL ACUITY WITH INTENSITY

M. H. PIRENNE DrSc. Ph.D.

Physiology Department, Marischal College
University of Aberdeen

1 The absolute threshold of vision
2 Experiments with large test fields
3 A retinal rod can be excited by one single quantum
4 Retinal threshold exceeds one quantum
5 Spectral sensitivity of the rods
6 Rods, cones and quanta at higher intensity levels
7 Spatial summation in the dark-adapted retina
8 Spatial summation and acuity
9 Conclusion
References

1. The Absolute Threshold of Vision

The minimum amount of light necessary for human vision is measured using a test field whose diameter subtends a few minutes of arc at the eye and which is illuminated in short flashes, each of which lasts, for instance, 0.001 sec. (Hecht, Shlaer & Pirenne, 1942). By using a fixation point, it can be arranged that this small field is imaged in the periphery of the dark-adapted retina. Under such conditions, it is found that about 100 quanta of light of wavelength 0.51 μ. passing through the pupil of the eye will as a rule be sufficient to make the human subject just see a dim, greyish, almost point-like, flash of light " out of the corner of his eye ". Part of this small quantity of light is lost by absorption in the eye media before reaching the retina, whilst the layer of visual purple contained in the retinal rods absorbs only a fraction of the light incident upon it. The actual stimulus acting upon the photoreceptors therefore consists of only a few quanta absorbed by the sensitive pigment contained in these receptors. Various considerations, to which we will return, prove that, in rod vision at any rate, the absorption of one quantum is not sufficient to make a man see a flash of light. Thus the " retinal threshold ", that is, the minimum number n of quanta which must be absorbed to give a sensation of light, is greater than one quantum.

There is no general agreement about the exact value of n. The light losses in the eye media and the absorption by the visual purple in the rods of the living human eye are not accurately known. Various indirect methods based on probability theory are also used for estimating the value of n. These methods suffer from uncertainties of their own ; there is no universal agreement about the interpretation of the experiments upon which they are based, or even about the experimental results themselves. Brumberg, Vavilov & Zverdlov (1943) suggest that n has values ranging from 18 to 32 according to the subject of the experiment; they also found the values 47 and 8 in two extreme cases. Bouman & van der Velden (1947) state that n is always equal to 2 quanta. Baumgardt (1948, 1950) also believes that $n = 2$.[1] Hecht *et al.* (1942) found values of n ranging between 5 and 8. Other experiments (Pirenne, 1946) give values from 4 to 10. A discussion (Pirenne & Denton, 1951 ; Denton & Pirenne, 1951a) of the experiments of Bouman and van der Velden suggests that n is very unlikely to be as small as 2 and is probably equal to 4 or more quanta. It may be pointed out that there seems to be no reason a priori why n should always have the same value, particularly when going from rods to cones, and from one organism to another.

Differences of opinion concerning the exact value of the retinal threshold n in man do not affect in any way the important fact that n is a small number, of the order of about 10 or less quanta. This represents an extremely small amount of energy, since the energy of a quantum is 3.9×10^{-12} erg for $\lambda = 0.51\ \mu$. Because the stimulus acting directly upon the retinal photoreceptors consists of such a small number of quanta, and because a quantum is of course the smallest quantity of light which can be emitted or absorbed by matter, it follows that statistical fluctuations of a purely physical nature must occur in the actual strength of the stimulus. From one trial to another, even though all experimental conditions are kept unchanged, the flash will deliver varying numbers of quanta to the retina. It is not intended here to develop this point which has been discussed elsewhere (Hecht *et al.* 1942; Pirenne, 1951). Reference may be made to recent work by Garlick & Wright (1952) who published recordings of somewhat analogous fluctuations in the electrical output pulses of a photomultiplier, a purely physical detector of light; the fluctuations were found to be negligible for light-pulses of high energy content, but became large when the average number of quanta in each pulse was small. On the other hand, a number of considerations argue in favour of a marked functional stability in the retina and visual system when they are working under the best conditions (Denton & Pirenne, 1952b). This gives further support to the view that the considerable variability of response which is observed near the threshold of vision is due for the greater part, at least in the best-controlled experiments, to physical fluctuations rather than to biological variations of the organism's sensitivity. The work of Hartline, Milne & Wagman (1947) on the fluctuation of response of single visual cells of the king crab *Limulus* will be of great importance in the present connexion, when it is published in full.

2. Experiments with Large Test Fields

Instead of small test fields and brief exposures, we may use large fields, for instance 50° in diameter or larger, emitting light continuously. A uniform overcast sky on a moonless night gives an idea of such a field. No special fixation point is used in this case. The field covers so wide an angle that, on account of ordinary eye movements, its image is likely to fall in turn on all parts of the retina, including the most sensitive regions in the periphery. In practice the exposures are not truly continuous but last, for instance, from 5 to 10 seconds. During such an exposure the total amount of light is much larger than the energy content of one of the flashes mentioned above, but here the interest of the experiment resides in the

[1] In a recent paper (*Naturwissenschaften*, 1952, 39, 388), Baumgardt suggests that, whereas 2 quanta are sufficient to cause an " elementary excitation ", a varying number of such " excitations " are necessary to produce a sensation of light. This apparently means that, on Baumgardt's theory, the number " n " *as defined above* is often greater than 2, being equal for instance to 4, 6 or 8. (Note added in proof.)

very low value of the retinal illumination rather than in the value of the total amount of light involved. When $\lambda = 0.51\ \mu$., the retinal illumination corresponding to the absolute threshold is of the order of 30,000 quanta sec.$^{-1}$ cm.$^{-2}$ (Denton & Pirenne, 1951b, 1952b). This is an upper value, assuming that there are no losses in the eye media. The number of rods in the retinal periphery is of the order of 10^7 rods/cm.^{2} and the diameter of the cross-section of the finer rods is about 2 μ. (Østerberg, 1935; see Pirenne, 1948b). The mean illumination per rod is thus about 0.001 quantum/sec. Assuming now that the probability of quantum absorption in the visual purple of the rods is 0.1, the mean amount of light absorbed per rod is 0.0001 quantum/sec. Since of course fractions of quanta cannot be absorbed, this means that, on the average, any given rod will absorb one quantum every 10,000 sec., that is, roughly every three hours. This result suggests that retinal photoreceptors cannot undergo a high rate of spontaneous excitation of the same nature as the excitation caused by light quanta. Otherwise this spontaneous background of "noise" would make it impossible for the nervous system to detect the very weak light "signal" which corresponds to the threshold stimulus (Denton & Pirenne, 1952b).

3. A Retinal Rod Can Be Excited by One Single Quantum

At the threshold of human vision, as we have seen, any given rod has about one chance in 100,000, or at the very most one in 10,000, of absorbing one quantum during any given period of 0.1 sec. One-tenth of a second is the order of the retinal action time (Graham & Margaria, 1935), the time during which physiological summation of stimulus can take place. It follows that the chance of any one rod absorbing two or more quanta within the action time is vanishingly small. If one quantum absorbed is insufficient to make the subject see light, summation must occur between the excitation processes initiated by single-quantum absorptions in different retinal rods.

According to Einstein's equivalence law, the first stage of a photochemical reaction consists of a number of individual processes in each of which one light-quantum reacts with one photosensitive molecule only. Studies in vitro show that the quantum yield in the bleaching of visual purple is independent of the wavelength over a wide part of the spectrum, and that it is probably unity, which means that there are no secondary reactions leading to further molecules of visual purple being bleached as the result of the initial absorption of one quantum (Schneider, Goodeve & Lythgoe, 1939; see Hecht, 1942). In the present case, therefore, the quantum absorbed in one rod will react on one molecule of visual purple only. This molecule must then initiate in the rod a trigger reaction leading to a process of nervous excitation capable of summation with other similar processes at the level of the bipolar cells or of the ganglion cells. The nature of this trigger reaction is not known. It must eventually bring into play thousands of ions or molecules, since a single nerve impulse involves such numbers of particles. As a reactive system, a retinal rod thus reaches the absolute limit set by quantum and molecular theory, since it can respond to the absorption of one single quantum acting upon one single molecule of sensitive substance.

4. Retinal Threshold Exceeds One Quantum

Light-quanta, it has been said, fall from the night sky like drops of rain. Each of these quanta, if absorbed by one element of the rod layer of the retina, can initiate in it a nervous reaction. But such an isolated reaction is not sufficient to make us see. It is probable that the excitation it causes in the retina does not pass the barrier of the bipolar or ganglion cell synapses. Accordingly we cannot detect individual quanta dropping from the sky. What we detect are small groups of quanta absorbed by the retina within sufficiently short intervals with regard to both time and space.

If one single quantum absorbed by the retina were capable of sending to the central nervous system a stimulation sufficient to make us see, the threshold of vision would—assuming that the retinal region considered is homogeneous in its properties—correspond to a certain mean total quantity of light, independently of the manner in which this nominal quantity is distributed over the retinal region. Let us suppose that a brief flash is used, and consider any two retinal areas A and B of similar properties. If one quantum absorbed is sufficient to make us see a flash of light, it is immaterial whether it is absorbed in area A or B. Now for a given mean, nominal energy content of the light flash, the chance of at least one quantum being absorbed by the retina will be the same when all the light is concentrated on A as when the same mean, nominal amount of light is divided between A and B. The probability of seeing therefore will be the same in the two cases. This can be mathematically demonstrated (Pirenne, 1951). If more than one quantum is needed for seeing, however, it is easy to see that the above conclusion cannot be generally true, unless complete physiological summation occurs over the whole retina. Assume for instance that six quanta are required and that summation takes place within A or within B taken separately, but not between A and B taken together. Then there will be a greater chance of six quanta being able to summate their actions when all the light is directed to A alone (or to B alone) than when it is shared between A and B.

Experiments (Denton & Pirenne, 1952a) suggested by this line of reasoning prove that the retinal threshold n is greater than one quantum, whilst at the same time giving indications about the extent of true spatial summation in the dark-adapted human retina. Two test fields A and B, each 10′ of arc in diameter, were placed at various distances from each other and from a fixation point. Green light of dominant wavelength 0.52 μ. could be flashed for durations of 0.0026 sec. on A alone, on B alone, or on A and B together. When the centres of fields A and B were 30′ apart and 20° from the fixation point, a typical result was the following: for A alone the frequency of seeing was $9/60 = 0.15$; for B alone, $6/60 = 0.10$; for A and B together, $23/60 = 0.38$ (the mean content of one flash per field being 36 quanta at the cornea). On the assumption that A and B act on independent retinal units, probability theory predicts that the frequency of seeing for A and B together would be $1 - ((1 - 0.15) \times (1 - 0.10)) = 0.24$. The experimental result for $A + B$ is higher than this, thus indicating the existence of physiological summation. This summation, however, is not complete. When the mean content of the flashes (72 quanta per field) was doubled, the frequency for A alone was 0.46 and for B alone 0.52; these values are higher than that for $A + B$ with 36 quanta per field. Such incomplete summation is at variance with what would be expected if n were equal to one quantum.

A result definitely disproving that $n = 1$ was obtained when A and B were placed 3.5° apart. Then A and B appeared to act independently of one another, a typical result being: A.

0.28; *B*, 0.30; *A* + *B*, 0.52 (102 quanta per field). The same calculation as above gives 0.50 for *A* + *B*, assuming they act independently. It is known from many similar experiments that doubling the mean content of the flashes (204 quanta per field) would have increased the frequency of seeing for *A* alone or *B* alone to about 0.80.

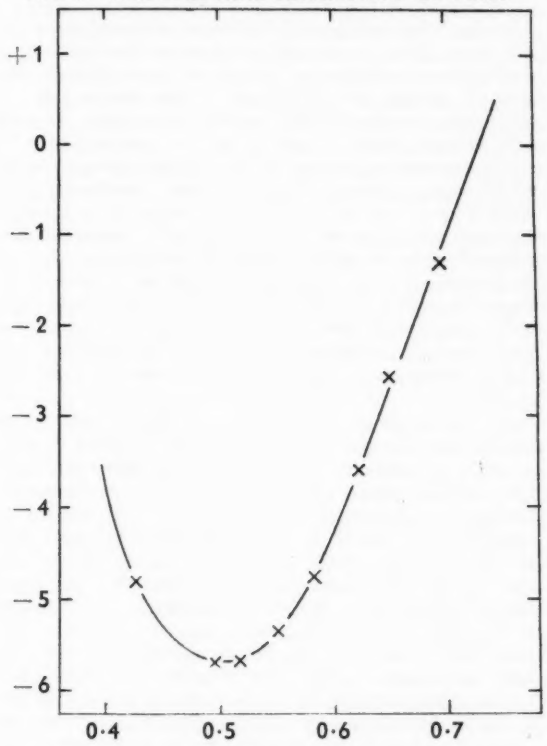
In discussing these experiments, it must be borne in mind that near threshold a flash of constant nominal intensity is sometimes seen and sometimes not seen, even under the best experimental conditions. This empirical fact of observation can be explained to a large extent on the basis of the physical fluctuations mentioned in the introduction, but the validity of the present reasoning is independent of this theoretical interpretation. Now it might appear puzzling that when both *A* and *B*, placed 3.5° apart, are used, the frequency of seeing at least one of them is greater than when *A* alone or *B* alone is used. But this finding is explained quantitatively by the variability of threshold response and does not imply any process of physiological summation between *A* and *B*. The following analogy may perhaps be used: if I make a bet in a horse race, and a second similar bet in another race, my chances of winning in at least one of the two races are greater than if I make a single bet; and my expenses are also greater. Similarly, in the present experiments, the average expenditure of light is doubled when *A* and *B* are used together, and the chances of seeing at least one of the two test fields increase. This increase is merely that to be expected if the parts of the retina receiving the image of the two fields respond independently of one another, as would be the case if they belonged to two different organisms. There is neither summation nor inhibition between them. A similar result was obtained in experiments on binocular absolute threshold (Pirenne, 1943). A more complete theoretical discussion (Pirenne, 1951) shows that such results prove that the retinal threshold *n* is greater than one quantum; it must be two or several quanta.

5. Spectral Sensitivity of the Rods

When light of wavelength shorter or longer than 0.51 μ. is used, a greater flux of energy must be emitted by unit area of the test field to reach the threshold of vision. Figure 1 shows the dependence of the logarithm of this threshold energy upon wavelength. Visual sensitivity may be defined as the reciprocal of threshold energy, so that log (sensitivity) = −log (threshold energy) and the curve of fig. 1 by a mere change of sign represents the sensitivity of the eye. The shape of the curve is the same as that of the scotopic visibility function determined by Crawford (1949) for 50 subjects by a matching method at a brightness level equal to 15 times the absolute threshold. Figure 1 shows that at the threshold the retinal illumination in ergs per second is about 1,000,000 times higher for 0.74 μ. than for 0.51 μ. The energy of a quantum is $h\nu$, where h is Planck's constant and ν the frequency of the light; this energy is therefore inversely proportional to the wavelength λ. When the retinal illumination is expressed in terms of the number of quanta per sec. instead of in ergs per sec., the values at the threshold for wavelengths 0.74 μ. and 0.51 μ. are in the ratio $(0.74/0.51)\times 10^6$ to 1. The actual value for 0.74 μ. is therefore $(0.74/0.51)\times 10^6\times 3\times 10^4$ or approximately 5×10^{10} quanta sec.$^{-1}$ cm.$^{-2}$ or about 1,500 quanta per sec. per retinal rod.

Now it is generally accepted, at least as a first approximation, that the number of quanta which must be absorbed by the rods in order to determine a given level of stimulation is the same whatever the wavelength of the light. Hence the probability of quantum absorption by the visual purple in the rods would be about 1,500,000 times smaller for red light of 0.74 μ. than for blue-green light of 0.51 μ., and thus the mean number of quanta absorbed in a given rod per second would be the same in both cases, i.e., 0.0001 quantum. For wavelengths different from 0.51 μ., the increased quantities of light which must strike the retina in order to stimulate it merely compensate the decreased chance of the light being absorbed by the photosensitive pigment in the rods. Thus the absorption curve of visual purple would have the same shape as the retinal sensitivity curve expressed on a quantum basis and corrected for absorption in the eye media. This is indeed verified for the middle of the visible spectrum, which gives strong confirmation of the theory. The theory may require some modification at the violet end and possibly also at the red end of the spectrum (Stiles, 1948; St George, 1952). It is not possible in vitro to measure light absorptions as low as some of those suggested by visual sensitivity measurements; technical difficulties have so far made it impossible to verify the theory in the far red of the visible spectrum.

FIG. 1. THE SPECTRAL SENSITIVITY OF MAN



Abscissae: wavelength (μ.)
Ordinates: log (threshold energy) (ergs cm.$^{-2}$ sec.$^{-1}$)

The points show measurements made as explained in text, section 5 (Denton & Pirenne, 1951b): the curve represents the relative scotopic visibility function (Crawford, 1949) suitably positioned by displacement parallel to the ordinate axis to show agreement with the points.

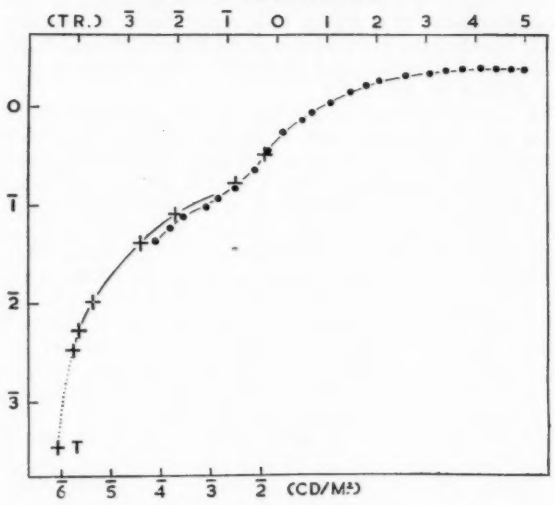
6. Rods, Cones and Quanta at Higher Intensity Levels

On the preceding basis two different physical lights will have equal effects on the rods and appear of the same brightness if the same number of quanta are absorbed from both lights by the layer of visual purple. This explains why rod vision does not discriminate between wavelengths as such. If the scotopic visibility curve of fig. 1 be used as a basis for calculation, and use be made of the fact that photometric and energy units are related by the so-called " mechanical equivalent of light " (Walsh, 1926), the amount of 0.51 μ. radiation equivalent to a given amount of white light of known spectral composition can be found. Now, in photometric units, the absolute threshold for white light corresponds to a field luminance of the order of 10^{-6} candela per square metre (Le Grand, 1948). It is found that this is indeed equivalent to the flux of radiation of wavelength 0.51 μ. which is emitted by the large field in the experiments of fig. 1, namely, 2×10^{-6} erg $sec.^{-1}$ $cm.^{-2}$.

Even on the darkest moonless nights, the sky often looks brighter than the ground and can be distinguished from it, the absolute threshold of vision for large fields being only rarely reached under natural conditions in the open air. Consider now a luminance of 0.01 cd $m.^{-2}$. This may be the luminance of a moonlit sky, thousands of times higher than that of a very dark night sky. Form vision, which is very poor in a very dark night, acquires considerable accuracy at such an intensity level (fig. 2). In objects silhouetted against the sky, details can then be seen which are only about eight times larger than the finest details visible in daylight (in objects of the same shape). For instance, a 1 mm. gap in a Landolt broken ring placed 1.14 m. from the eye can be seen. The best acuity at this level is mediated by the parafovea, and is probably due to the co-operation of rods and cones (Shlaer, Smith & Chase, 1942). Now the luminance of 0.01 cd $m.^{-2}$ corresponds to 10,000 times the absolute threshold and, according to the above approximate calculations, to an average of one quantum absorbed in a given rod per second, or rather less because of a greater pupillary contraction. It is estimated (Broda, Goodeve & Lythgoe, 1940 ; see Hecht, 1942) that the terminal segment of a rod of the dark-adapted frog retina contains about 10^8 molecules of visual purple. A human rod is smaller; let us assume it contains about a million visual purple molecules. The quantum absorbed will act upon only one such molecule, at least in the primary stage. As has been emphasized by Baumgardt (1950), it is clear that at this level of light intensity, and even at higher levels, we are dealing with chemical reactions of a very special kind. At such light intensities, the purely photochemical side of light-adaptation, due only to the partial bleaching of visual pigments, is probably unimportant, since the number of quanta absorbed is so small in comparison with the total number of pigment molecules available in the photoreceptors.

If the foveal cones are linked to the brain by individual optic nerve fibres, and if no summation occurs between one cone and another, it is obvious that each cone will need to absorb at the very least one quantum if each cone unit is to be excited. Neglecting the difference in shape, size and spectral sensitivity between cones and rods, and assuming for the sake of argument that the probability of quantum absorption is as high in a cone as in a rod, each cone would on the average also absorb one quantum per second at a luminance of 0.01 cd $m.^{-2}$. Since the physiological summation

FIG. 2. VISUAL ACUITY VERSUS INTENSITY, USING THE WHOLE RETINA



Acuity and intensity are both plotted on logarithmic scales. The first branch of the curve corresponds mostly to rod vision, the second mostly to cone vision. Visual acuity is expressed as the reciprocal of the angle, in minutes, subtended at the eye by the gap of the smallest Landolt C which can just be resolved.

Ordinates: log acuity

Lower abscissae: luminance, in candelas per square metre (CD/M.²), of the field against which the black test objects were presented

Higher abscissae: retinal illumination, in trolands (TR.) (Le Grand, 1948, p. 107). One troland (0 on the log scale) corresponds to about 20 quanta of the most effective wavelength striking a retinal area 2 μ. in diameter, every second

+ unpublished measurements made in 1942-45 in Professor Sir F. Bartlett's Laboratory, Cambridge, on behalf of the Military Personnel Research Committee

T absolute threshold for a 50° field

● measurements by Shlaer (1937) on his subject A.M.C.

time is shorter than one second, the single-cone units cannot all be brought into activity at an intensity of 0.01 cd $m.^{-2}$. Yet we know that cones begin to come into play at this level.

At intensities of the order of 1 cd $m.^{-2}$, corresponding to twilight or dim daylight, we must expect each cone to absorb only a few quanta during each period of physiological summation. Thus even when foveal acuity is about unity, that is, nearly half the maximum value it can reach at the highest brightness, it appears unlikely that the absolute threshold of the least sensitive among the retinal photoreceptors is greatly exceeded. It is not even certain that it is always reached.

7. Spatial Summation in the Dark-Adapted Retina

If we return to the experiments made with two light spots *A* and *B* placed 0.5° apart, the results show that a certain degree of true physiological summation exists between the corresponding retinal areas. That is, the excitations from *A* and from *B* are able to help each other, but the resulting frequency of seeing remains less than if the light were con-

centrated on to *A* or *B* alone instead of being divided between *A* and *B*. The retinal regions corresponding to *A* and *B* do not respond independently of one another; thus partial physiological summation can take place in the periphery of the retina across an unstimulated gap corresponding to a visual angle of about 20′ of arc (i.e., 0.5° less an allowance of 10′ for the finite sizes of *A* and *B*). We have seen that, when the unstimulated gap is 3° 20′, there is neither summation nor inhibition. Thus it seems that the periphery of the dark-adapted human retina contains independent functional units at least 0.5°, but less than 3.5°, across, summation taking place to a certain extent inside each of these units.

These experiments may be compared to those in which the absolute threshold is measured for retinal areas of increasing size. Experiments of the latter kind indicate that in the periphery, for areas of diameter up to 30′ or 1°, the total amount of light received at threshold by the retinal area remains almost constant; physiological summation is almost complete. For larger areas the total amount of light necessary increases (Le Grand, 1948; Weinstein & Arnulf, 1946). There is a general measure of agreement between these results and those of the experiments made with two discrete areas *A* and *B*, but in the latter experiment the degree of summation seems to be less.

Electrophysiological studies have revealed that optic nerve fibres of the frog have large receptive fields of the order of 0.5 mm. in diameter in the retina, that these receptive fields overlap one another, and that spatial summation occurs in the centre of the field (Hartline, 1940). Giant ganglion cells with wide receptive fields have also been found in the retina of a vertebrate, namely, the cat (Rushton, 1949, 1950).[2] The above experiments on summation, as well as a number of other considerations, suggest that the dark-adapted human retina also behaves under certain conditions as if it contained such units having a diameter of the order of 1° or 2°. A circular area subtending a visual angle of 2° contains about 40,000 rods and has a diameter of 0.58 mm. on the retina. (The theory of Bouman & van der Velden (1947), based on $n = 2$, assumes that the peripheral retina is made up of units all of which are about 10′ in diameter. This assumption seems hardly compatible with the preceding evidence.)

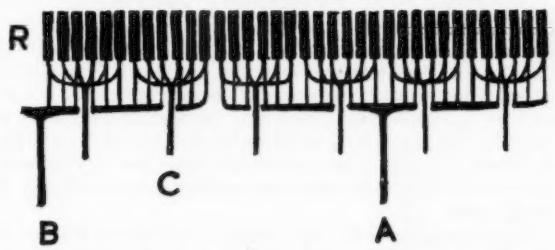
8. Spatial Summation and Acuity

At suprathreshold levels, visual acuity in the periphery reaches values corresponding to visual angles of the order of 10′, which is markedly smaller than the angle of 1° corresponding to the large units just discussed. It seems probable therefore that the retina also contains functional units smaller than these. Anatomical studies show the existence of bipolar cells and of ganglion cells covering fields of various sizes in the mammalian retina (Polyak, 1941). It would be an economical arrangement if the photoreceptors were common to both large and small units, the smaller units coming into play when the light intensity is higher. The following imaginary scheme may illustrate this idea (fig. 3). Consider 40,000 neighbouring rods connected to one ganglion cell with single optic nerve fibre. These rods are subdivided into 100 groups of 400 rods, each group being again connected to a separate ganglion cell and optic nerve fibre. In order to stimulate either the large unit or any one of the small units, a minimum of about six quanta must be absorbed in the corresponding receptive field during a retinal action time of 0.1 sec. When the mean total number of quanta absorbed by the large unit is, for instance, ten quanta, the threshold of this unit will be passed and the unit will be continuously excited. But at the same time the smaller units receive only an average of 0.1 quantum and therefore remain generally inactive. If the intensity is increased 100 times, however, the threshold of the small units will be reached. Thus, without any switching mechanism to change nervous connexions in the retina, a finer mosaic of retinal units will come into play when the intensity is increased.

At intensities near the absolute threshold for large fields only the largest retinal units can come into play. Further, there must be a lower range of intensities in which even they are not all excited, since there is an intensity range over which it is uncertain whether a retinal unit will respond or not (Pirenne, 1948a). The numbers of units excited will increase with intensity until an intensity is reached at which the probability of exciting any of the units is unity. At that stage all the units are excited. As the recruiting of the units progresses with increasing intensity the unstimulated gaps between excited units will become smaller and eventually disappear. The retinal mosaic formed by the large units will therefore become functionally finer and visual acuity will increase with increasing intensity. The results of experiments on the visibility of black discs presented 10° or 20° from the fovea in the centre of a wide field exposed in flashes agree with this theory (see fig. 4) (Pirenne, 1946). A fixation point was used in these experiments. The acuity-intensity curves obtained have a first branch ending as a small plateau at acuity value of 1/96 or 1/120. Then a break occurs and the acuity rises again. Special measurements in the region of transition showed that the break is genuine. It has not yet been possible to pursue the experiments at higher intensities but it was found that the relative position of the two branches was unaffected when blue light was used instead of white, which proves that the second branch is due like the first to rod vision, not cone vision. It will also be noted that the breaks in the curves of fig. 4 occur at acuity levels at least 10 times lower than the rod-cone break of fig. 2.

If the kind of anatomical and functional arrangement suggested above exists in the retina, we should expect the same kind of recruiting to occur at higher intensities among the smaller rod units. Thus as intensity increases, starting from the level of the absolute threshold for large fields, there

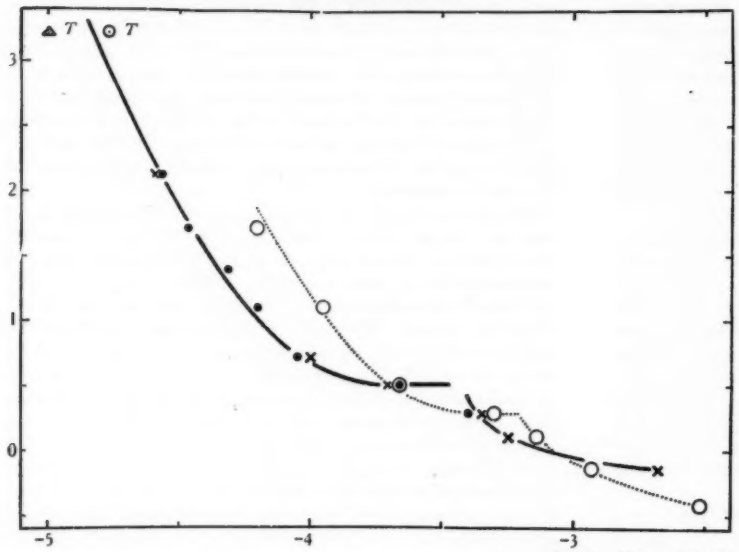
FIG. 3. DIAGRAM OF THE TYPE OF RETINAL CONNEXIONS SUGGESTED IN THE TEXT



A large functional unit, A, and part of another, B, are shown. The same receptors, R, which form A and B, are also connected in a different way to form smaller units such as C. At sufficiently high intensities the small units absorb enough quanta to be able to respond but at low intensities they do not, and only large units summate enough light to remain active.

[2] See also Rushton, p. 68 of this Bulletin.

FIG. 4. DETECTION OF BLACK OBJECTS BY THE PERIPHERY: WITH FIXATION



Reproduced by permission of the Cambridge Philosophical Society

The test object was a black disc presented in the middle of a 45° field which could be illuminated in 0.03-second flashes. The field brightness necessary to the dark-adapted subject for detecting the presence of the black disc when the light was flashed was determined for various sizes of disc.

Abscissae: log brightness of screen, in equivalent foot candles
The abscissae of the points T give two determinations of the absolute threshold of the large field itself.

Ordinates: logarithm of the area (in square degrees) of the disc, which is on the limit of detection

× ● } results obtained 20° from the fovea

○ results obtained 10° from the fovea

It should be noted that the quantity plotted is the area of the disc; the smaller this area, the greater the acuity. The first branch of the curve for 10° thus reaches a higher acuity than that for 20°; it is also shifted towards higher intensities. This agrees with the theory that the functional units involved are smaller and less sensitive in terms of illumination, when going from the periphery towards the fovea (from Pirenne, 1946).

would first be a recruiting of the largest retinal units, very sensitive in terms of retinal illumination, but forming a very coarse mosaic giving only rudimentary vision of form. Then, at higher intensities, this would be superseded by a finer mosaic, more accurate for form vision but made up of units smaller and less sensitive in terms of retinal illumination. More than two such steps may of course exist in rod vision, and at a certain level cones may co-operate with rods in retinal units, particularly in the parafovea (Shlaer *et al.* 1942). In foveal vision it is also conceivable that cones are similarly arranged in functional units, the smallest units consisting of single cones.

Concerning the retina as a whole, it is well known that in the natural use of the eye the region giving the most accurate vision of details shifts from the periphery to the parafovea and then to the fovea as light intensity increases. We have seen that at intensities where a fairly high foveal acuity is reached the absolute threshold of single cones, if reached, cannot be greatly exceeded. Thus by going from one region of the retina to another, there occurs here a kind of recruiting similar to that which has just been postulated to explain the intensity-acuity relationship in one and the same region of the periphery. (The curve of fig. 2 shows the two branches corresponding to rod and to cone vision, but inside the two main branches there are no breaks such as have been observed in the above experiments using fixation. This is most probably due to the fact that the retinal region changed progressively, the subject always using the best part of his retina at each intensity.)

9. Conclusion

Further experiments will be required to investigate the detailed implications of the theory outlined here. The fundamental evidence for the theory comes from the absolute values of the retinal illuminations corresponding to various levels of visual acuity, taken in conjunction with the obvious fact that the absolute threshold of a retinal unit cannot be lower than one absorbed quantum. The theory states that, over a considerable part of the intensity-acuity range, the increase of acuity is due to the bringing into play of finer and finer retinal mosaics made up of smaller and smaller functional units, the *absolute* thresholds of the various units being reached successively as the retinal illumination increases. Complications may be expected, notably on account of the probable presence of "off" fibres (Hartline, 1940) in the human retina. The saccadic movements, which make the eye take discrete "snapshots" with exposures of the same order as the retinal action time, may on the other hand make it easier to elucidate the normal functioning of the visual system on the present basis.

The present theory, which was developed in collaboration with Dr E. J. Denton, takes up theories and suggestions from Hecht (Hecht, 1928; Hecht & Wald, 1934), von Buddenbrock (1937) and Lythgoe (1940). On account of the approximate nature of some of the calculations given here, it can hardly be decided at the present stage whether the intensities for which the theory may hold reach the plateau where foveal acuity is maximal. In any case, beyond a certain level of intensity, it must be expected that the absolute threshold of all the retinal units is passed. Recruiting of retinal units might then still occur with regard to certain suprathreshold levels of excitation, instead of the absolute threshold level.

ACKNOWLEDGEMENTS

I am greatly indebted to Dr W. S. Stiles for many discussions on the above problems and for editorial advice in preparing this paper. I also wish to thank Professor A. V. Hill, for his encouragement and unfailing help over many years, Sir Frederic Bartlett, in whose laboratory experiments leading to the present theory were initiated, Sir Henry Dale, O.M., Sir Eric Rideal, Sir Charles Lovatt Evans, and Professor E. W. H. Cruickshank, in whose laboratories they were continued, and the Medical Research Council for their support of these investigations.

REFERENCES

Baumgardt, E. L. M. (1948) *J. gen. Physiol.* **31,** 269

Baumgardt, E. L. M. (1950) *Les théories photochimiques classiques et quantiques de la vision et l'inhibition nerveuse en vision liminaire.* Revue d'Optique, Paris

Bouman, M. A. & Velden, H. A. van der (1947) *J. opt. Soc. Amer.* **37,** 908

Broda, E. E., Goodeve, C. F. & Lythgoe, R. J. (1940) *J. Physiol.* **98,** 397

Brumberg, E. M., Vavilov, S. I. & Zverdlov, Z. M. (1943) *J. Phys., Moscow,* **7,** 1

Buddenbrock, W. von (1937) *Grundriss der vergleichenden Physiologie,* 2nd ed. Borntraeger, Berlin

Crawford, B. H. (1949) *Proc. phys. Soc. Lond.* B, **62,** 321

Denton, E. J. & Pirenne, M. H. (1951a) *J. Physiol.* **115,** 68*P*

Denton, E. J. & Pirenne, M. H. (1951b) *J. Physiol.* **115,** 66*P*

Denton, E. J. & Pirenne, M. H. (1952a) *J. Physiol.* **116,** 32*P*

Denton, E. J. & Pirenne, M. H. (1952b) *J. Physiol.* **117,** 55*P*

Garlick, G. F. J. & Wright, G. T. (1952) *Proc. phys. Soc. Lond.* B, **65,** 415

Graham, C. A. & Margaria, R. (1935) *Amer. J. Physiol.* **113,** 302

Hartline, H. K. (1940) *Amer. J. Physiol.* **130,** 690

Hartline, H. K., Milne, L. J. & Wagman, I. H. (1947) *Fed. Proc.* **6,** 124 [Abstract]

Hecht, S. (1928) *J. gen. Physiol.* **11,** 255

Hecht, S. (1942) *Ann. Rev. Biochem.* **11,** 465

Hecht, S., Shlaer, S. & Pirenne, M. H. (1942) *J. gen. Physiol.* **25,** 819

Hecht, S. & Wald, G. (1934) *J. gen. Physiol.* **17,** 517

Le Grand, Y. (1948) *Optique physiologique.* Vol. 2: *Lumière et couleurs.* Revue d'Optique, Paris

Lythgoe, R. J. (1940) *Brit. J. Ophthal.* **24,** 40

Østerberg, G. (1935) *Acta Ophthal., Kbh.* Suppl. No. 6

Pirenne, M. H. (1943) *Nature, Lond.* **152,** 698

Pirenne, M. H. (1946) *Proc. Camb. phil. Soc.* **42,** 78

Pirenne, M. H. (1948a) *J. Physiol.* **107,** 47*P*

Pirenne, M. H. (1948b) *Vision and the eye.** Chapman & Hall, London

Pirenne, M. H. (1951) In: Butler, J. A. V. & Randall, J. T., ed. *Progress in biophysics and biophysical chemistry,* vol. 2, p. 193. Pergamon Press, London

Pirenne, M. H. & Denton, E. J. (1951) *J. opt. Soc. Amer.* **41,** 426

Polyak, S. L. (1941) *The retina: the anatomy and the histology of the retina in man, ape, and monkey, including the consideration of visual functions, the history of physiological optics, and the histological laboratory technique.* University Press, Chicago, Ill.

Rushton, W. A. H. (1949) *Nature, Lond.* **164,** 743

Rushton, W. A. H. (1950) *J. Physiol.* **111,** 26*P*

St George, R. C. C. (1952) *J. gen. Physiol.* **35,** 495

Schneider, E. E., Goodeve, C. F. & Lythgoe, R. J. (1939) *Proc. roy. Soc.* A, **170,** 102

Shlaer, S. (1937) *J. gen. Physiol.* **21,** 165

Shlaer, S., Smith, E. L. & Chase, A. M. (1942) *J. gen. Physiol.* **25,** 553

Stiles, W. S. (1948) In: *Transactions of the Optical Convention of the Worshipful Company of Spectacle Makers,* p. 97. Spectacle Makers Company, London

Walsh, J. W. T. (1926) *Photometry.* Constable, London

Weinstein, C. & Arnulf, A. (1946) *Comm. lab. Inst. Opt.* **2,** 1. Cited by Le Grand (1948)

* Reviewed on p. 83 of this Bulletin.

* * *

ELECTRIC RECORDS FROM THE VERTEBRATE OPTIC NERVE

W. A. H. RUSHTON Sc.D. F.R.S.

Physiological Laboratory
University of Cambridge

1 Identification of Granit's retinal " elements "
2 Excitatory and inhibitory fields of single fibres
3 Records from single fibres in the optic nerve trunk
4 Significance of the optic nerve record
References

Before valve amplification became available to physiologists, the mode of action of sense organs could be studied only by intelligent appreciation of their structure and by the analysis of motor responses in animals and of sensory impressions in man—responses which were usually several steps removed from the action of the sense organ studied. The magnificent work done in spite of these limitations by such masters as Thomas Young, Hermann von Helmholtz, and many others, is common knowledge, but only within the last 30 years has it been possible to get an accurate idea of the actual message signalled by the sense organ to the central nervous system.

Pioneer and pre-eminent in this field is Adrian, who in his *Basis of sensation* in 1928 had already studied the way in which the principal sense organs act and was able to generalize the nature of their sensory messages in nerves—generalizations which have well stood the test of subsequent critical study and development from every centre of neurophysiology in the world.

In 1927 and 1928 Adrian and Matthews made their well-known researches upon the discharge from a vertebrate optic nerve. A conger eel was chosen because of its anatomical suitability, and amplified electrical records from the whole optic nerve showed that this portion of the central nervous system conducted impulses of the same general pattern as in peripheral nerves, but that the relation between stimulus and record was greatly complicated by the synaptic network of the retina. For instance, it was found that, though on histological grounds each optic nerve fibre must on average connect with a patch of retina at least 90 μ. in diameter, the area actually connected was nearer 1 mm. in diameter and thus there is a very great overlap in the receptive fields of individual nerve fibres (Adrian & Matthews, 1927a, 1927b, 1928).

The matter was taken to a new level of precision ten years later when Hartline (1938, 1940) succeeded in extending to fish, frogs and reptiles his beautiful studies upon single optic nerve fibres, which hitherto had been confined to invertebrates, chiefly the arachnoid *Limulus*. He dissected small bundles of fibres from the frog's optic disc and, placing these upon fine electrodes, obtained in favourable cases records from single fibres. He established that the optic nerve discharges that had been known to arise when the light was turned " on " or " off " do not necessarily occur in the same nerve fibres, for some respond only to " on " (20% in the frog) and some (30%) only to " off ". He also investigated the size of the receptive field of a single optic nerve fibre and confirmed the estimate made by Adrian and Matthews of a patch nearly 1 mm. in diameter. The summation of light stimuli within these overlapping fields was also investigated.

About the same time Granit began the series of researches upon single retinal " elements " which has covered so large a range of vertebrate types and so many aspects of temporal and chromatic stimulation (Granit & Svaetichin, 1939). Granit's technique was simpler than Hartline's: instead of isolating the fibre by dissection, the retina was left in situ, only the front of the eye being opened and the lens removed. There was then applied to the retina a microelectrode consisting of a 25 μ. platinum wire, sealed into a glass capillary tube so that the wire tip just protruded. The tip was lowered gently on to the retina and moved over it by a micromanipulator, and it was found possible to obtain in this way electric records clearly derived from single excitable units—Granit's " elements ". A disadvantage of Granit's method is that it is not quite certain what is being recorded. But the great advantage is that the eye can be studied with far less operative disturbance than by Hartline's method, and innumerable records have been obtained from the eyes of mammals and birds, as well as from the cold-blooded eyes studied by Hartline. It is true that when a microelectrode is simply placed upon a suitable element in the living retina, the investigation might terminate at any minute by a sudden flick of the eye. Yet, in a decerebrate cat under a light urethane anaesthesia, Granit has often obtained a sensitive retina with perfect immobility for several hours, and little ambiguity is ever introduced through uncontrolled eye movements.

The results of Granit and his school are too varied and extensive for any summary here, nor is this necessary as they have been brought together in full historical perspective in his own well-known monograph (Granit, 1947). The question of what in fact is recorded by Granit's microelectrode, however, will be taken up, since some objections have been raised to the belief that the spikes represent optic nerve discharges. The matter has been investigated by the present writer and leads to conclusions essentially identical with those previously adopted by Granit. This work was done upon the cat in Professor Granit's laboratory in 1949 and has appeared only in summary form (Rushton, 1949, 1950).

1. Identification of Granit's Retinal " Elements "

It is obvious from inspecting Granit's published records that they consist of responses of one (or occasionally more) single excitable " elements " which discharge an action potential in the usual all-or-none fashion with a duration of 1–2 msec. A single medullated nerve fibre might behave like this, but there are objections to supposing that the electrode was actually picking up the responses of single optic nerve fibres as they coursed over the retina, much as Hartline picked up the responses of his dissected fibres. The optic nerve fibres are very fine (mostly 1 μ. or 2 μ., though they may perhaps occasionally reach 15 μ.) and they tend to run together in bundles. A 25 μ. wire would not be expected to record large isolated spikes of 300 μV. from such a structure, and Thomson, recording from the optic nerve itself (see section 3).

has never got much more than 50 μV. Nor, if by some fortunate application of the electrode this were achieved, would we expect its stability to be maintained over a period of minutes or even hours in the presence of small respiratory and other movements. Moreover, no large spikes can be obtained from an electrode applied to the optic disc itself, though a mass of fine small spikes is easily seen. Thus, in appreciating Granit's records, we are faced with the following questions.

Why are the spikes so large and independent of minute shifts of the electrode ? Why are they found only in certain spots of the retina ? Are they due to some typical structure which is at that spot especially well placed for recording? Or is the structure a rare one and, if so, is it pathological? Or are the large spikes due to the synchronous firing of a patch of retinal cells? Adrian & Matthews (1928) had found evidence of synchronous discharges over a whole region of retina and it is known that there are various conditions which favour this behaviour. It is true that such discharges are apt to present a spontaneous rhythm quite different from most of Granit's records, but a somewhat different kind of synchrony, possibly presided over by an amacrine cell, might well account for the facts.

The first investigation was to verify the impression, obtained by moving the microelectrode over the retina in the dark, that a large spike could still be obtained when the microelectrode was moved 50 μ. or more in any direction. Since the eye was open and flaccid and the vitreous viscid, it was possible that this movement was deceptive, and that the electrode dragged the retina about with it in any direction for a short way. This possibility was excluded by two methods.

i. A little " rope " of seven strands of fine insulated silver wire was made and varnished to form a rod. When this was cleanly cut across, a cluster of seven microelectrodes was obtained. The distance between the centres of adjacent wires was 80 μ. (electrode 3 was defective). Each wire could be connected separately to the amplifier through a multipolar switch. Figure 1 shows the records taken in rapid succession in the order 0 → 6 → 0 after the electrode system had been adjusted so that record 0 was as large as possible. It is seen that there is little sign of movement of retina relative to electrodes during the course of this series, since spikes remain about the same size in the left and right records. If all these spikes are from the same excitable structure, therefore, the extent of spread is more than 80 μ. in all directions. It can easily be shown that the spikes are in fact all from the same structure. For if, say, record 0 was from one element and record 1 from another, then by connecting wires from 0 and 1 simultaneously to the amplifier input we should obtain a record in which the spikes both of 0 and 1 would appear, but each would be halved in amplitude by the shunting of the other. In short, the combined record would show spikes at about twice the frequency and half the amplitude. This was never observed. It is easy to connect electrode 0 to 1 or any other during the course of a discharge, and the frequency was never modified, but the amplitude became about the average of that obtained in records 0 and 1 alone. This can only mean that the same element is recorded in each of the records of fig. 1 and that its field of electric response is extensive in all directions.

ii. A second method, which confirms this result, is to remove the vitreous from the eye and to observe, through a dissecting microscope, the actual movement of the microelectrode over

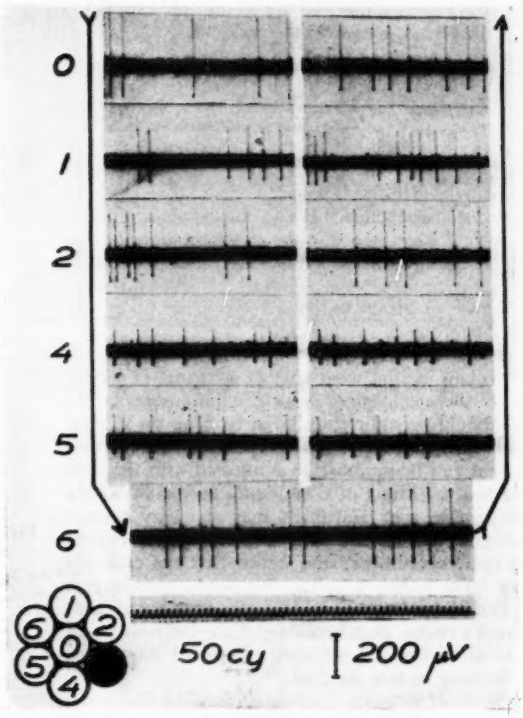


FIG. 1. Action potential spikes recorded from the retina of a cat. Records were made using 6 out of a cluster of 7 microelectrodes (microelectrode 3 was defective), arranged as shown in bottom left-hand corner. Records were taken in the order indicated by the arrows.

the brightly illuminated retina. The red vessels lying over the yellow-green tapetum form excellent landmarks. It was seen that the electrode could slip over the surface without moving the vessels and it was presumed that other retinal structures were also unmoved. Figure 2 shows a curve indicating the size of the spike of a single element as the microelectrode moved across the region of its maximum effect. The size is plotted upwards from the horizontal line, which also gives the space scale. Here again we see that some 100 μ. in either direction from the maximum are needed before the spike falls to half size (where it is largely obscured by the presence of new " elements ").

The next question is to find whether this rather large field of activity is due to one cell or a synchronous mass of cells. In these experiments, which were performed in collaboration with Dr K. O. Donner, the microelectrode was a very sharp needle insulated except for the extreme tip. The idea was to lower this needle on to the excitable element and damage it. If the element were one cell, the activity should cease abruptly, but a synchronous mass might be expected to break up into smaller groups or perhaps be relatively unaffected by puncture.

The results are shown in fig. 3. Records *a–e* show unequivocally a sudden stop in the discharge as the electrode was pressed steadily down, supporting the idea of a single cell.

Records *f* and *g* were of all our records the two which accorded least well with the idea; they are, nevertheless, consistent with it.

The mechanical effect of the puncture was to stimulate the element to a high frequency discharge which rose to a shriek in the loudspeaker. The fact that the spike size gradually diminished with very high frequency would be expected of any single nerve element. In record *g* alone is there some evidence of a step-like decline. It was not met on any other occasion. On the whole the evidence suggests that there may be single large excitable cells responsible for these spikes, but if histology revealed no such cells the results would not be irreconcilable with the synchrony of smaller cells.

Before we turn to the histological study, it is worth considering what sort of structure is needed to fill the electrophysiological requirements. The cell should be sufficiently large to have an electric field spreading 100 μ. in all directions. It responds to light and hence must be connected with the rods and cones. Many of Granit's results are so exactly comparable with Hartline's that one may suspect that Granit's impulse pattern is also the pattern of the optic nerve fibres and hence that the cells are very large ganglion cells. We might thus expect to find these large ganglion cells distributed over the cat's retina, each separated from the next by an interval of, say, 0.2–0.5 mm. Figure 2 shows this to be more or less the case.

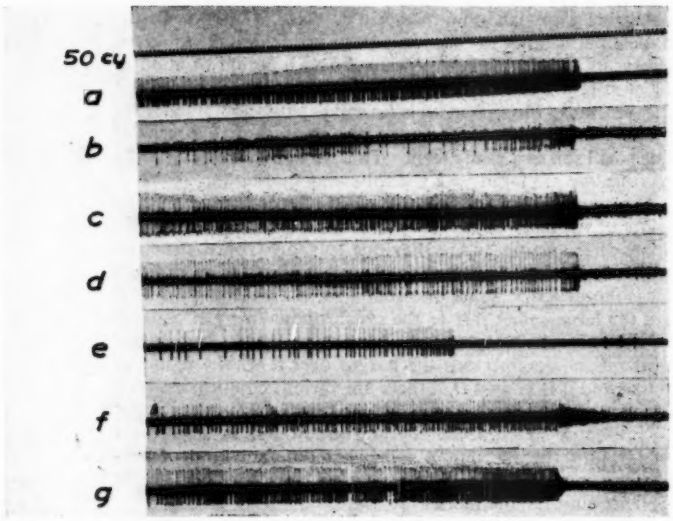


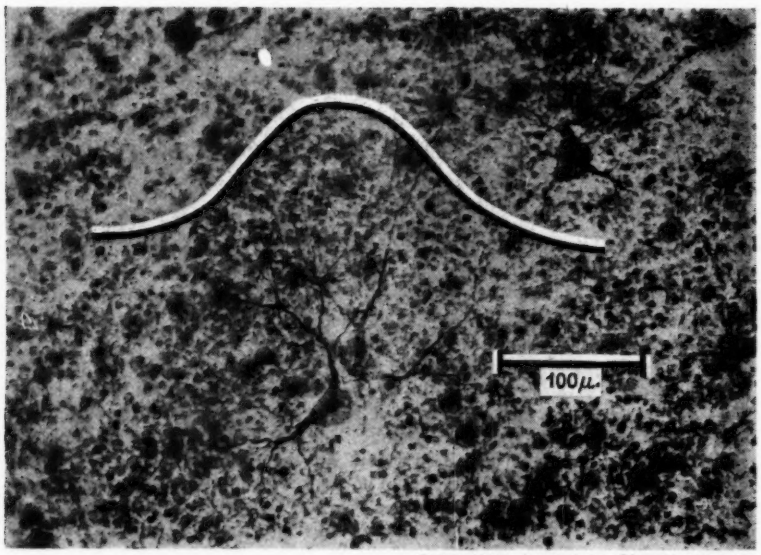
FIG. 3. Action potentials recorded from single " elements " in the retina of a cat. Microelectrode used was a sharp needle, which punctured the " element."

Records *a* to *e* show abrupt termination of discharge.

Record *f* shows a very high frequency terminal discharge, which ceased by gradually dying away; this often happened.

Record *g* shows a step-like decline; this was unique.

The most valuable method of staining was a modification of the silver method of Palmgren (1948, 1951). The retina was carefully dissected from the eye, fixed flat and stained and mounted whole (without sectioning). Figure 2 shows the presence of large cells with cell bodies of 30–50 μ. and branches which spread in all directions, and actually spread over a circle of 1 mm. diameter. Some of these cells were seen to be the cell bodies of fibres of the optic nerve. They were found to be distributed all over the retina, with a density of the order required to fit electrical expectations. If the cell and all its branches fired off synchronously, this might very well account for the space distribution of electric field observed (fig. 2) (though the recorded field distribution was not obtained from the particular cell shown).



Rushton (1950), by permission of Journal of Physiology

FIG. 2. Three giant ganglion cells in the retina of a cat. The curve (taken from a different eye) shows how the spike height varies as the microelectrode is moved straight across an " element ". The 100 μ. line gives the space scale and forms the zero for spike height.

Finally, an attempt was made to correlate definitely these large cells with Granit's " elements ". The vitreous was removed from the eye and a camera lucida drawing made of the main blood vessels as they lay over the tapetum. Now the surface was explored with the microelectrode and each place where a spike was best obtained was marked upon the map by indicating the image of the tip of the microelectrode. In this way a map of " elements " in relation to the retinal vessels was made. Two days later, after fixing and staining, the map was correlated with the mounted whole retina. This correlation was surprisingly hard to do, for the map represented one kind of distortion from the original spherical retina and the final flat section a quite different one. However, in many cases the topology could be recognized with certainty (by

arterial branching) and a fair correlation found between "elements" and these large cells. There were nearly always more cells in a region than discharging "elements", but it is known that by no means all "elements" will discharge in the bright steady light used in this investigation.

Because of the difficulties in obtaining simple general correlations, one places greater weight upon exceptional multiple correlations such as in the results of fig. 4. The action potential (fig. 4a) showed clearly two independent elements, one of which appeared to be more distant because it was smaller. But changing the position of the electrode did not change the relative size of these two spikes; they varied in fixed proportion no matter what was the direction of movement. This clearly suggested that the two "elements" were of different sizes and situated very close together. When, two days later, the site of this record was studied on the stained retina, the map showed it to be at the edge of the blood vessel, as shown by the circle (fig. 4b). The way that the optic nerve fibres had been displaced where they lay on the vessel indicated that at some time the electrode had indeed pushed its way upwards (and perhaps had displaced the cells as well as the fibres). What is striking is that there are in fact two nerve cells of different sizes extremely close together. In addition to the cells shown, there was one at the zero of the scale, and no other within a radius of 0.5 mm. from the circle.

The conclusions from this investigation appear to support Granit's interpretations and to overcome the difficulties which seemed to stand in the way. For if the "elements" are large ganglion cells, the impulse patterns in them are the impulse patterns of the optic nerves attached. On the other hand, the difficulties mentioned earlier in accepting the records as being led from single retinal fibres no longer apply.

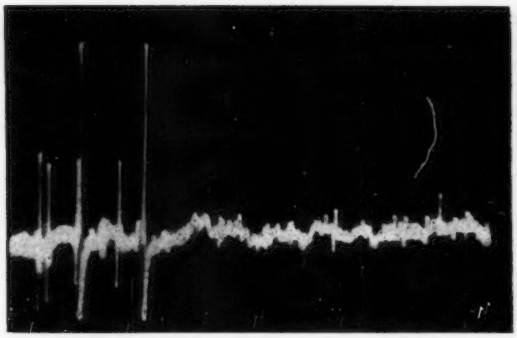
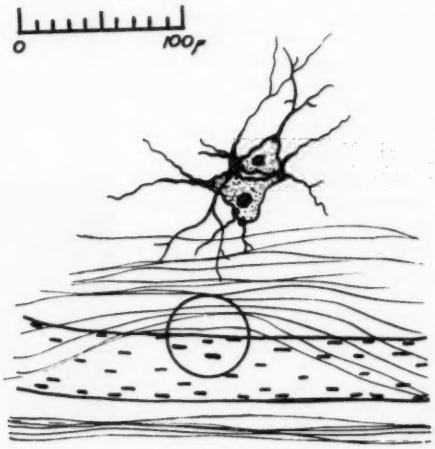


FIG. 4a. Action potential record from the retina of a cat, showing discharge of two "elements", one twice the size of the other. The relative sizes of the spikes could not be changed by moving the electrode.



Rushton (1949), by permission of Nature

FIG. 4b. Camera lucida tracing of stained retina in the region where the spikes shown in fig. 4a were obtained. The circle marks the places at the edge of the artery where the electrode was. Note how it has pushed aside the nerve fibres from on top of the vessel and how two different sized ganglion cells are close together in the vicinity.

2. Excitatory and Inhibitory Fields of Single Fibres

An extension of Hartline's observations on the frog by the use of Granit's type of electrode has recently been made by Barlow (1950) working in Cambridge. A 20 μ. platinum wire in a plastic insulator of refractive index close to that of water cast hardly any shadows on the retina, and permitted investigations of the receptive fields of ganglions by exploration with a small point of light.

The interpretation of the structures recorded in this way in the frog is in the main similar to what has been described above for the cat. Two kinds of response could be obtained, as Gernandt (1948) found in the cat. The high-pitched crackle, hard to isolate and small in amplitude, was certainly tapped from the optic nerve fibres as they coursed over the frog's retina. For, as is seen in fig. 5, this type of response is obtained (inset 2) when the electrode lies on the line between the spot illuminated and the optic disc. The larger isolated "plop" sound (inset 1) is obtained only when the illuminated spot is close to the electrode and is believed to be the discharge of a ganglion cell. Such ganglion spikes can be recorded from almost any region of the retina and hence do not represent so isolated a structure as the giant ganglion cells in the cat.

By this technique many of Hartline's results were confirmed. The "on" fibres were found in more or less the stated proportions. The size of the receptive field was something less than 1 mm.^{2}, and the ganglion cell was sensitive both to summation and to movement of light within this region. The nature of the summation was studied in greater detail, and resulted in an important new discovery.

A small spot of light was moved about in the receptive field and in each position its threshold intensity was found. Plotting the reciprocal of this intensity gave the sensitivity map of the receptive field for the fibre. After this, the spot was placed in the centre of the receptive field and increased in size. The larger the spot, the lower the threshold, and therefore the higher the sensitivity. It was found that this increased sensitivity was fairly exactly the sum of the sensitivities of all the points on the map upon which the (uniform) light was falling. It would thus be expected that when the limits of the receptive field were reached, no further lowering of threshold would occur on further increasing the

size of the spot. In the case of the "off" fibres, this was in fact so. But with the "on-off" fibres a new relationship appeared. For in this case the increase in the size of the spot beyond the edge of the receptive field caused a fall of sensitivity. Now the large spot is physically identical with a central spot just occupying the receptive area of the fibre plus a ring of light surrounding this. When the ring alone is shone there is no effect. When the central spot alone is shone there is a good discharge. When the spot plus the ring are shone together the discharge is less than with the spot alone. Thus the ring outside the receptive field inhibits the discharge from the centre.

This is shown more directly in fig. 6, which gives the results when two spots of light are employed. One of these (centre spot) is in the middle of the receptive field of the ganglion cell below the microelectrode. When it alone is turned on and off, the ganglion discharges are as shown in the upper two frames. The second spot (side spot) is situated just outside the receptive field and consequently has no effect when it alone is turned on or off. But as may be seen from the lower two frames, when the two spots are turned on and off together, the discharge from the centre spot is much inhibited. There is a linear relation between the exciting action of the central spot and the inhibiting action of the peripheral, and (for the "off" response) about 100 units of peripheral light are needed to neutralize one central unit.

The sharpening of contours by simultaneous contrast has often been ascribed to inhibition, though not necessarily to inhibition in the retina itself. Its demonstration in this simple and precise form in the excised eye is the beginning of exact understanding of the phenomenon. It appears that some important work in this field has recently been done in Baltimore by Kuffler. His contribution to the Cold Spring Harbor Symposium in 1952 is not yet available, but at the risk of some inaccuracy it seems worth indicating here the very striking observations which I understand that he has made.

Kuffler has used Granit's type of electrode, but introduced it behind the limbus in the unopened eye of the cat, so that the animal's own optical system could be used, with virtually normal circulatory conditions. He has found that the size of the receptive field for a single ganglion cell is of the order of 1 mm.^{2} but that it depends upon the brightness of the background illumination, being greatest in the dark. This confirms in a precise way the conclusion of Craik & Vernon (1941) that in dark-adaptation the human "summation area" increases, and shows that at least part of this summation is retinal.

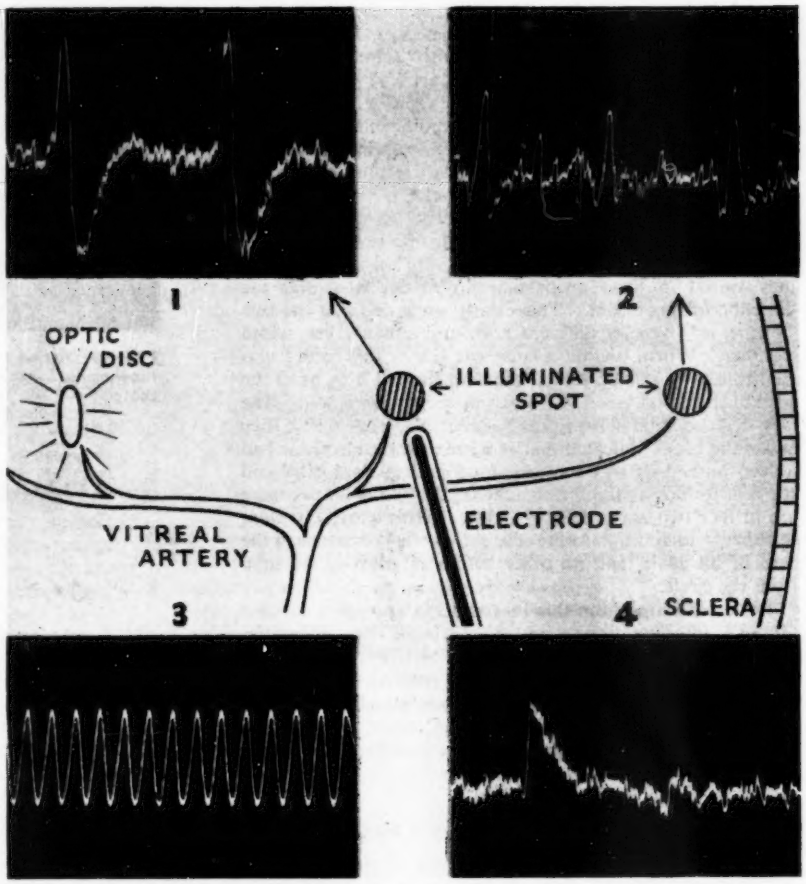


FIG. 5. Diagrammatic drawing of retina and electrode showing position of stimulus, and records of responses.

1: record of slow type of action potential, obtained when electrode is close to illuminated spot.
2: record of fast type of action potential, obtained when electrode lies on the line between illuminated spot and optic disc.
3: record of pulse at 1,000 cycles per second.
4: record of 20 μV. negative pulse.

Another important observation of Kuffler's is that "on" or "off" discharges are not entirely fixed by the ganglion cell but depend too upon the place on the receptive field excited by the small light spot used. If the centre of the field is "on" then the periphery will be "off" and vice versa. With suitable employment of two spots of light, the one may be made to inhibit the effect of the other. The publication of the details of these observations will undoubtedly constitute a valuable advance in our knowledge of the organization of the retina.

3. Records from Single Fibres in the Optic Nerve Trunk

This account may end, where it started, with the technique of recording from the optic nerve outside the eye. This has the advantage that the eye need not be disturbed in any way

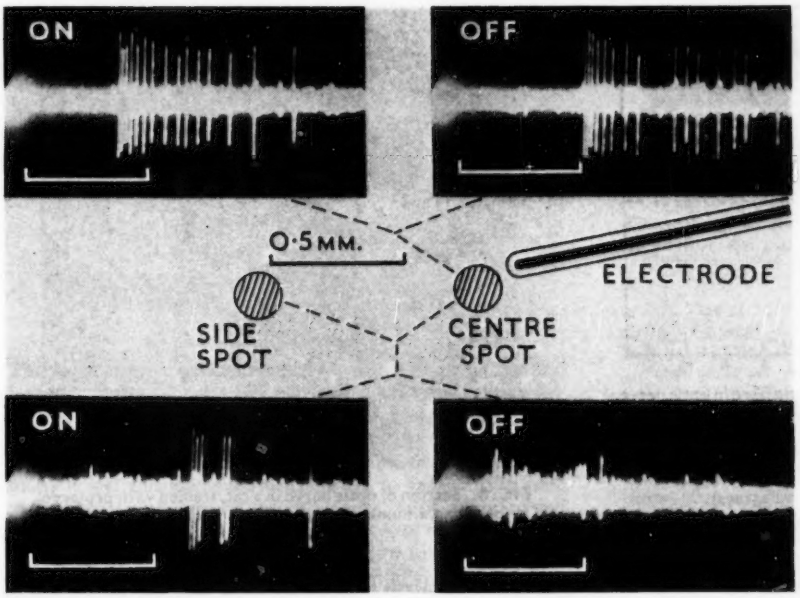


FIG. 6. Diagrammatic drawing and records of responses produced by stimulation near the centre of the receptive field of a ganglion cell, showing inhibiting effect of illuminating a side spot.

Top records : action potentials with illumination of centre spot on, and off, no illumination of the side spot.

Bottom records: action potentials with illumination of both centre and side spots on, and off, together.

Time unit shown: 0.2 sec.

nor is there any shadow of the electrode to disturb the retinal image. The problem of how to obtain records from single fibres in the optic nerve has been solved by L. C. Thomson in London by the construction of a specially fine metal microelectrode. The details of its construction are important and will soon be published, but in principle a fine glass capillary with a 10 μ. tip is filled with metallic silver by electrolysis, so that a crystal projects some 5 μ. beyond the tip. In this way an all-metal contact is made from nerve to amplifier. As is seen in fig. 7, a single large spike (50 μV.) is well enough distinguished from background noise and smaller spikes for its study to be free from ambiguity. And the curve below shows the threshold intensity required to elicit this spike when the light spot (33 μ. in diameter on the retina) was shone at various angular distances from the centre of the receptive fields. Each degree corresponds to 100 μ. of retinal distance. The spikes are positive-going in their first (and larger) phase, and Thomson accepts the probable explanation that the microelectrode has penetrated and made contact with the axis cylinder of the recorded fibre.

Analysis with the use of this technique has only just begun, but it seemed worthy of mention since the method of recording has some features superior to all other known methods.

4. Significance of the Optic Nerve Record

In one way the records so far obtained from the vertebrate optic nerve seem unsatisfactory. In the case of touch receptors or stretch receptors or pain endings, the response patterns in the nerves correspond well enough to our sensory impressions. It is not hard to understand in a general way how these impulse trains may pass with minor modifications to the central nuclei and there be correlated with what we perceive. But at present the relation between visual sensation and the impulse patterns which have been recorded from optic nerve fibres is quite bewildering. One is forced to ask: Are the records we have typical of an average optic nerve? Are the sensations we experience derived from an average optic nerve? Can we argue from responses in frogs and cats, to sensation in man?

Investigators have probably always picked the largest ganglion cells and nerve fibres they could find because the record is then easier to obtain, to maintain stable, and to interpret against background disturbance. We have seen (fig. 2) that in the cat there is a giant ganglion cell which is probably responsible for most of the nerve patterns hitherto obtained. Cajal and Polyak have both observed that large cells connect with large optic nerve fibres, and Professor W. E. Le Gros Clark has sent me a section of cat's optic nerve (fig. 8) which shows large nerve fibres in a proportion which might very well correspond to the giant cells. Thomson's large spikes (fig. 7) were obtained by his lowering the microelectrode and neglecting small spikes until a 50 μV. spike was encountered. So, in all this work on the mammal, a very large integrating ganglion cell appears to have been selected for analysis. We can hardly conclude without more information that this is typical of all mammalian optic nerve fibres.

In the case of the frog the situation is rather different. Barlow claims that a really good electrode will record from almost any place on the retina a ganglion response of one or other of Hartline's types, and it is unlikely that Hartline, with his small bundles of optic nerve fibres, would have failed to notice if there had been some small fibre discharges of a type quite different from those he described. One might conclude then that all the nerves of the frog discharge according to Hartline's types, and that the large mammalian nerves do roughly likewise. Whether the smaller mammalian nerves have quite a different response pattern is at present unknown.

Turning to the question of our own visual sensations, it would appear from introspection that these are dominated by foveal vision, no doubt because the fovea has so large a cortical representation. But if it is the pattern of impulses in the *foveal* nerves that is the basis of ordinary visual sensation, this will clearly not be typical of an average optic nerve fibre and still less is it likely to correspond with the discharge from giant ganglion cells. Indeed, it is unlikely that records from any animal below the primate can be considered comparable to the human foveal pattern—unless it be *Limulus*!

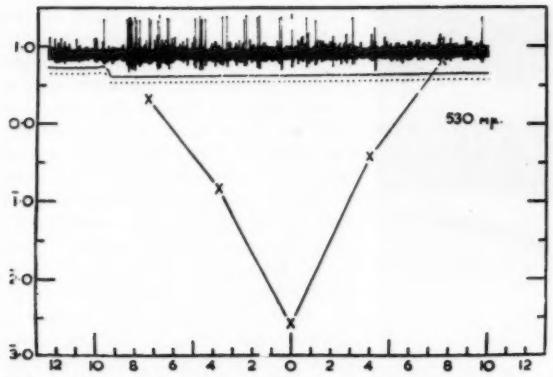


FIG. 7. Record of action potentials from single fibre in optic nerve, outside the eye, showing 50 μV. spikes. Curve shows the logarithm of the threshold intensity required to elicit this spike when light (of wavelength 530 mμ.) fell at various angular distances from centre of receptive field.

Abscissae: angular distances (in degrees)
Ordinates: log threshold energy

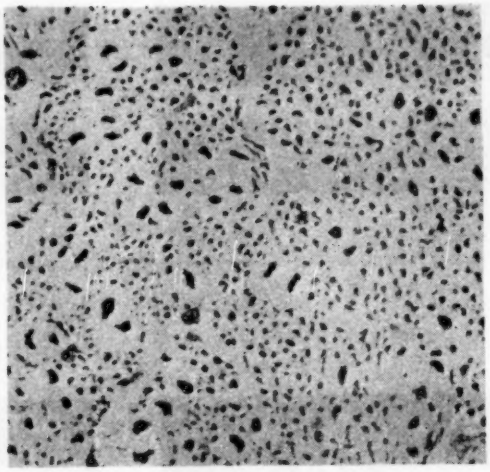


FIG. 8. Section of optic nerve in a cat, stained with protargol, showing a number of large fibres (magnification × 500).

In eyes generally, the enormous reduction in pathways from photoreceptors to optic nerves makes it certain that most of the detail received at the receptor level is sacrificed for economy in transmission. As Adrian has put it, the information received is drastically edited and only the more dramatic features signalled to the brain. In the primate fovea, however, Cajal and Polyak have shown that an exception is made and, by a system of private lines through the midget bipolars, nearly all the photoreceptor information may be transmitted to the cortex where it becomes anatomically possible for a much fuller analysis to be made. Now since this system will be really effective only if no gross editing has taken place on the way, we should expect that the cortex would receive a message very much like the discharge of the foveal cones themselves.

This then may well account for the curious fact that if we wish to find an optic nerve record which seems to correspond to our sensations, we should turn not to the frog or to the cat, where the record has been edited by giant ganglion cells, but to the simple responses from *Limulus* which come straight from photoreceptors.

REFERENCES

Adrian, E. D. (1928) *Basis of sensation: the action of the sense organs.* Christophers, London
Adrian, E. D. & Matthews, R. (1927a) *J. Physiol.* **63,** 378
Adrian, E. D. & Matthews, R. (1927b) *J. Physiol.* **64,** 279
Adrian, E. D. & Matthews, R. (1928) *J. Physiol.* **65,** 273
Barlow, H. B. (1950) In: *XVIII International Physiological Congress, Copenhagen, August 15–August 18, 1950. Abstracts of communications,* p. 88 [Copenhagen]
Craik, K. J. W. & Vernon, M. D. (1941) *Brit. J. Psychol.* **32,** 62
Gernandt, B. (1948) *Acta physiol. scand.* **15,** 88
Granit, R. (1947) *Sensory mechanisms of the retina: with an appendix on electroretinography.* Cumberlege, London
Granit, R. & Svaetichin, G. (1939) *Upsala LäkFören. Förh.* **65,** 161
Hartline, H. K. (1938) *Amer. J. Physiol.* **121,** 400
Hartline, H. K. (1940) *Amer. J. Physiol.* **130,** 690; 700
Palmgren, A. (1948) *Acta zool., Stockh.* **29,** 377
Palmgren, A. (1951) *Acta zool., Stockh.* **32,** 1
Rushton, W. A. H. (1949) *Nature, Lond.* **164,** 743
Rushton, W. A. H. (1950) *J. Physiol.* **111,** 26*P*

* * *

PHYSIOLOGY OF VISION

A Bibliography of British Books and Periodicals

KATHARINE TANSLEY D.Sc.

Wernher Research Fellow (Medical Research Council)
Institute of Ophthalmology, London

The physiology of vision is a subject which is of interest and importance to many workers, not only to professional physiologists but also to those concerned with visual defects, with illumination, with industrial health and experimental psychology. The problem of colour vision in particular penetrates many fields, and interested workers are to be found in such industries as photography, optics, dyeing, paint, ceramics and textiles, as well as in fields where good colour vision is important, such as some branches of the armed forces and of the railway services. In Great Britain much of the best contemporary work on colour vision is, at present, being done in physics laboratories.

There is no specialist journal devoted to the physiology of vision and most work is, therefore, published either in clinical journals or in journals concerned with physiology, biochemistry, general biology or general science. In addition, papers are to be found in journals of experimental psychology and illuminating engineering.

A number of books is available dealing with different aspects of vision, although none covers the whole field. There is still no substitute in any language for Helmholtz's famous *Physiological optics*. In Britain a new interest in vision began about 1925 when Lythgoe, Stiles and Wright all independently began work on the subject. For this reason the arbitrary date 1924 has been taken as a limit for the present bibliography, although one or two slightly earlier publications have been included. Between 1926 and 1935 the Physiology of Vision Committee of the Medical Research Council published a series of reports, most of them giving the results of original investigations, one or two of them classics in their field and all but one now, unfortunately, out of print.

Purely ophthalmological books or those dealing with clinical techniques have not been included, since they can always be found in the *Quarterly cumulative index medicus*.

A. PERIODICALS

The data given are name of publication, frequency of appearance now, place of publication, publisher and amount of annual subscription.

1. Specialized Journals

a. CLINICAL

British Journal of Ophthalmology. Monthly. London: British Medical Association. 84s.
This journal can be purchased together with *Ophthalmic Literature* for an over-all subscription of 147s. About half the papers are not clinical.

British Orthoptic Journal. Annually. Shrewsbury: Wilding. 10s. 5d.

Ophthalmic Literature. Seven times yearly. London: British Medical Association. 84s.
An abstracts journal mainly covering the clinical literature but with sections on anatomy, embryology, heredity, physiology and biochemistry of the eye, physiology and psychology of vision and illumination. Review articles are published once a year of which about one in four is not clinical.

Proceedings of the Royal Society of Medicine. (Ophthalmological Section.) Monthly. London: Lewis. 126s.
Papers read before the Ophthalmological Section of the Royal Society of Medicine. Mainly clinical.
The Ophthalmological Section does not appear in every number of the Proceedings.

Transactions of the Ophthalmological Society of the United Kingdom. Annually. London: Churchill. 40s.
Papers read before the Ophthalmological Society of the United Kingdom. Mainly clinical.

b. OPTICAL

British Journal of Physiological Optics. Three times yearly. London: British Optical Association. 42s.
Some purely theoretical papers on general problems.

c. ILLUMINATION

Light and Lighting. Monthly. London: Illuminating Engineering Society. 20s.

Transactions of the Illuminating Engineering Society. Ten times yearly. London: Illuminating Engineering Society. 42s.
Papers read before the Illuminating Engineering Society.

2. Physiological and Anatomical Journals

Brain. Quarterly. London: Macmillan. 50s.
Occasional papers reporting experimental work or new observations on the retina and its central nervous connexions.

Journal of Anatomy. Quarterly. London: Cambridge University Press. 63s.
Occasional papers reporting new observations on the anatomy and embryology of the eye.

Journal of Physiology. Monthly. London: Cambridge University Press. 210s.
Papers reporting original experimental work, and abstracts of papers read before the Physiological Society. Much of the fundamental work on the physiology of vision has been and is still published in this journal.

Quarterly Journal of Experimental Physiology. Quarterly. London: Griffin. 50s.
Very occasional papers on the eye.

3. Biochemical, Pharmacological and Nutrition Journals

Biochemical Journal. Monthly. London: Cambridge University Press. 168s.
Occasional papers reporting experimental work on biochemical problems related to the eye, and abstracts of papers read before the Biochemical Society.

British Journal of Nutrition. Quarterly. London: Cambridge University Press. 60s.
Very occasional papers dealing with the effects of deficiency diseases on the eye and on vision. Also abstracts of papers read before the Nutrition Society.

British Journal of Pharmacology and Chemotherapy. Quarterly. London: British Medical Association. 84s.
Occasional papers reporting experimental work on the effect of drugs on ocular structures.

4. Psychological Journals

British Journal of Psychology (General Section). Quarterly. London: Cambridge University Press. 40s.
Some papers dealing with the psychology of vision.

Journal of Mental Science. Quarterly. London: Churchill. 50s.
Occasional papers, mainly psychiatric, on visual problems.

Journal of Neurology, Neurosurgery and Psychiatry. Quarterly. London: British Medical Association. 42s.
Occasional papers dealing with the neurological and psychiatric aspects of vision.

Quarterly Journal of Experimental Psychology. Quarterly. Cambridge: Heffer. 30s.
Some papers reporting experimental work on the psychology of vision.

5. Zoological and Genetic Journals

Annals of Eugenics. Quarterly. London: Cambridge University Press. 57s.
Very occasional papers on the inheritance of visual abnormalities in man.

Journal of Genetics. Irregular. London: Cambridge University Press. 50s. per volume.
Very occasional papers dealing with the inheritance of eye abnormalities in man and animals.

Journal of the Linnean Society of London (Zoological Section). Irregular. London: Longmans Green. Price variable.
Very occasional papers on the morphology of the visual apparatus in different animals.

Journal of the Marine Biological Association of the United Kingdom. Irregular. London: Cambridge University Press. Price variable.
Very occasional papers on the visual reactions of marine invertebrates.

Proceedings of the Linnean Society of London. Irregular. London: Longmans Green. Price variable.
Papers read before the Linnean Society of London. Very occasional papers on the physiology of vision in animals.

Proceedings of the Zoological Society of London. Quarterly. London: Zoological Society of London. 80s.
Papers read before the Zoological Society of London. Very occasional papers on the eye and vision in animals.

6. Journals of General Biology

Biological Reviews of the Cambridge Philosophical Society. Quarterly. London: Cambridge University Press. 60s.
Occasional review articles on the comparative physiology and anatomy of the eye.

Journal of Experimental Biology. Quarterly. London: Cambridge University Press. 63s.
Occasional papers on the comparative physiology and biochemistry of vision. Papers on the visual behaviour of invertebrates can often be found in this journal.

New Biology. Irregular. Harmondsworth: Penguin Books. 9s. for four issues (including *Science News, q.v.*).
Occasional general articles of a semi-popular nature.

Philosophical Transactions of the Royal Society of London. (B. Biological Sciences.) Irregular. London: Cambridge University Press. Price according to length.
Papers which usually report original work, but which are too long for publication in the usual scientific journals. Very occasionally there is one on the eye or on vision.

Proceedings of the Royal Society of London. (B. Biological Sciences.) Irregular. London: Cambridge University Press. 57s.
Papers read before the Royal Society of London, a few of which report original work on the physiology of vision. Much of the fundamental work, especially on colour vision, has been published in this journal.

7. Physical Journals

Proceedings of the Physical Society. Six times yearly. London: The Physical Society. 42s.
Papers read before the Physical Society and its groups, including the annual Thomas Young Oration. Several of the fundamental papers on vision, especially on colour vision, have been published here. Papers read before the Colour Group of the Physical Society, many of which deal with the physiology of vision, are published in these Proceedings.

Proceedings of the Royal Society of London. (A. Mathematical and Physical Sciences.) Irregular. London: Cambridge University Press. 47s. 6d. per volume.
Very occasional papers on the physical and chemical aspects of vision.

8. Journals of General Science

Discovery. Monthly. Norwich: Jarrold. 19s.
Occasional popular articles.

Nature. Weekly. London: Macmillan. 120s.
Short notes, in the form of letters to the editor, often containing important new observations. Occasional longer articles either theoretical or briefly reporting experimental results.

Science News. Irregular. Harmondsworth: Penguin Books. 9s. for four issues (including *New Biology, q.v.*).
Occasional general articles of a semi-popular nature.

B. BOOKS

Data given are author's name, title of book, edition, date and place of publication, publisher, price (if in print), whether illustrated or containing a bibliography. Unfortunately no certain information can be given as to availability, but it is likely that many books published before 1940 are out of print.

1. Physiology of Vision

Davson, H. *The physiology of the eye.* 1938, London: Churchill. 32s. Illustrated. Short bibliography.

Duke-Elder, W. S. *Text-book of ophthalmology.* Vol. 1. 1932, London: Kimpton. 63s. Illustrated. Bibliography for each section.
The best available full general account of the development, morphology and physiology of the visual system.

Granit, R. *Sensory mechanisms of the retina.* 1947, London: Cumberlege. 35s. Illustrated. Bibliography.
An account of twenty years' work on the electrical responses of the retina. This book contains the best discussion of visual purple and visual violet to be found anywhere, as well as the exposition of a hypothesis of colour vision which, although not generally accepted in its entirety, is now a classic. Professor Granit is in no sense a British worker although he spent three years at Oxford during which he obtained some of the results incorporated in this book.

Lythgoe, R. J. *Practical physiology of the sense organs.* 1934, Oxford: University Press; London: Milford. Illustrated. 1s.
Description of experiments on sense organ physiology suitable for the medical student.

Mann, I. & Pirie, A. *The science of seeing.* 1950, Harmondsworth: Penguin Books. 2s. 6d. Illustrated.
A popular account.

Medical Research Council. *The legibility of print.* By R. L. Pyke. (Medical Research Council Special Report Series No. 110.) 1926, London: His Majesty's Stationery Office. Illustrated. Bibliography.
Report of an original investigation.

Medical Research Council. *Reports of the Committee on the Physiology of Vision. I. Illumination and visual capacities.* By R. J. Lythgoe. (Medical Research Council Special Report Series No. 104.) 1926, London: His Majesty's Stationery Office. Bibliography.
A review of the literature.

Medical Research Council. *Reports of the Committee on the Physiology of Vision. II. Dark adaptation.* By D. R. Adams. (Medical Research Council Special Report Series No. 127.) 1929, London: His Majesty's Stationery Office. Bibliography.
A review of the literature.

Medical Research Council. *Reports of the Committee on the Physiology of Vision. III. Two studies in the psychology of reading.* By M. D. Vernon & R. W. Pickford. (Medical Research Council Special Report Series No. 130.) 1929, London: His Majesty's Stationery Office.
Report of two original investigations.

Medical Research Council. *Reports of the Committee on the Physiology of Vision. IV. Experiments on binocular vision.* By N. M. S. Langlands. (Medical Research Council Special Report Series No. 133.) 1929, London: His Majesty's Stationery Office. Illustrated. Bibliography.

Medical Research Council. *Reports of the Committee on the Physiology of Vision. V. The adaptation of the eye: its relation to the critical frequency of flicker.* By R. J. Lythgoe & K. Tansley. (Medical Research Council Special Report Series No. 134.) 1929, London: His Majesty's Stationery Office.
Report of an original investigation. This is still the classical publication on flicker.

Medical Research Council. *Reports of the Committee on the Physiology of Vision. VI. Some experiments on peripheral vision.* By M. Salaman. (Medical Research Council Special Report Series No. 136.) 1929, London: His Majesty's Stationery Office. Illustrated.

Medical Research Council. *Reports of the Committee on the Physiology of Vision. VII. A re-determination of the trichromatic mixture data.* By W. D. Wright. (Medical Research Council Special Report Series No. 139.) 1929, London: His Majesty's Stationery Office. Illustrated. Bibliography.

Medical Research Council. *Reports of the Committee on the Physiology of Vision. VIII. The movements of the eyes in reading.* By M. D. Vernon. (Medical Research Council Special Report Series No. 148.) 1930, London: His Majesty's Stationery Office. Bibliography.
Report of an original investigation.

Medical Research Council. *Reports of the Committee on the Physiology of Vision. IX. Psychological factors in peripheral vision.* By G. C. Grindley. (Medical Research Council Special Report Series No. 163.) 1931, London: His Majesty's Stationery Office. Bibliography.
Report of an original investigation.

Medical Research Council. *Reports of the Committee on the Physiology of Vision. X. The measurement of visual acuity.* By R. J. Lythgoe. (Medical Research Council Special Report Series No. 173.) 1932, London: His Majesty's Stationery Office. Bibliography.
Report of an original investigation and one of the classical publications on visual acuity as affected by adaptation.

Medical Research Council. *Reports of the Committee on the Physiology of Vision. XI. Individual differences in normal colour vision.* By W. O'D. Pierce. (Medical Research Council Special Report Series No. 181.) 1933, London: His Majesty's Stationery Office. Bibiliography.
A review of the literature.

Medical Research Council. *Reports of the Committee on the Physiology of Vision. XII. Colour vision requirements in the Royal Navy.* (Medical Research Council Special Report Series No. 185.) 1933, London: His Majesty's Stationery Office.

Medical Research Council. *Reports of the Committee on the Physiology of Vision. XIII. Determination of the sensitiveness of the eye to differences in the saturation of colours.* By L. C. Martin, F. L. Warburton & W. J. Morgan. (Medical Research Council Special Report Series No. 188.) 1933, London: His Majesty's Stationery Office.
Report of an original investigation.

Medical Research Council. *Reports of the Committee on the Physiology of Vision. XIV. Characteristics of dichromatic vision.* By F. H. G. Pitt. (Medical Research Council Special Report Series No. 200.) 1935, London: His Majesty's Stationery Office. 1s. 3d. Illustrated. Bibliography.
Report of an original investigation.

Physical & Optical Societies. *Report of a joint discussion on vision held on June 3rd, 1932, at the Imperial College of Science.* 1932, London: Physical Society.

Pirenne, M. H. *Vision and the eye.* 1948, London: Pilot Press. 12s. 6d. Illustrated. Bibliography for each chapter.
A physicist's view of the physiology of the retina.

Wright, W. D. *The perception of light.* 1938, London: Blackie. 7s.
An elementary account primarily for the physicist.

2. Colour Vision

Boltz, C. L. *A statue to Mr Trattles.* 1952, London: Butterworth. 12s. 6d. Illustrated.
Popular science essays of which two deal with normal and defective colour vision.

Conference on Colour Vision. Proceedings in: *Documenta ophthalmologica: advances in ophthalmology.* Vol. 3. 1949, The Hague: Junk. 80s.
Most of the papers read at the Conference on Colour Vision held in Cambridge, between 28 July and 2 August 1947, are published in this volume. Some of the contributions to the discussions are also included.

Hartridge, H. *Colours and how we see them.* 1949, London: Bell. 15s. Illustrated.
Christmas lectures at the Royal Institution.

Hartridge, H. *Recent advances in the physiology of vision.* 1950, London: Churchill. 25s. Illustrated. Bibliography.

Murray, H. D. *Colour in theory and practice.* 2nd ed., 1952, London: Chapman & Hall. 70s. Illustrated. Bibliography to each chapter.
This is a joint work by the editor and some others.

Parsons, J. H. *Colour vision.* 1924, Cambridge: University Press. Illustrated. Bibliography.
One of the classical publications on colour vision.

Physical Society. *Report on defective colour vision in industry.* 1946, London: Physical Society. 3s. 6d.
A report by a committee of the Colour Group of the Physical Society.

Pickford, R. W. *Individual differences in colour vision.* 1951, London: Routledge & Kegan Paul. 30s. Illustrated. Bibliography.
Report of a psychologist's investigations.

Willmer, E. N. *Retinal structure and colour vision.* 1946, London: Cambridge University Press. 21s. Illustrated. Bibliography.
Sets forth an unorthodox theory of colour vision which is not generally accepted.

Wright, W. D. *The measurement of colour.* 1944, London: Hilger. 30s. Illustrated. Bibliography.

Wright, W. D. *Researches on normal and defective colour vision.* 1946, London: Kimpton. 36s. Illustrated. Bibliography.
An account of nearly twenty-five years' research on colour vision.

3. Illumination in Industry

Campbell, D. A., Riddell, W. J. B. & MacNalty, A. S. *Eyes in industry.* 1951, London: Longmans Green. 30s. Illustrated. Bibliography.

Department of Scientific and Industrial Research. Illumination Research Committee. *The terminology of illumination and vision.* (Technical papers, No. 1.) 1926, London: His Majesty's Stationery Office.

Department of Scientific and Industrial Research. Illumination Research Committee. *Penetration of daylight and sunlight into buildings.* (Technical papers, No. 7.) 2nd ed., 1932, London: His Majesty's Stationery Office.

Department of Scientific and Industrial Research. Illumination Research Committee. *The effect of glare on the brightness difference threshold.* By W. S. Stiles. (Technical papers, No. 8.) 1929, London: His Majesty's Stationery Office. Illustrated.
Report of a laboratory investigation.

Department of Scientific and Industrial Research. Illumination Research Committee. *The effect of distribution and colour on the suitability of lighting for clerical work.* (Technical papers, No. 10.) 1930, London: His Majesty's Stationery Office. Illustrated.
A laboratory investigation.

Department of Scientific and Industrial Research. Illumination Research Committee. *Appraisal of street lighting installations.* By W. S. Stiles & C. Dunbar. (Technical papers, No. 13.) 1931, London: His Majesty's Stationery Office.
A field survey.

Department of Scientific and Industrial Research. Illumination Research Committee. *The appraisal of glare from motor car headlights.* By W. S. Stiles & C. Dunbar. (Technical papers, No. 16.) 1935, London: His Majesty's Stationery Office.

Department of Scientific and Industrial Research. Illumination Research Committee. *Seasonal variation of daylight illumination.* (Technical papers, No. 17.) Reprinted 1948, London: His Majesty's Stationery Office. 6d. Illustrated.
Report of an investigation.

Department of Scientific and Industrial Research. Road Research Committee. *Interim measures for the prevention of dazzle on roads.* (Technical papers, No. 14.) 1949, London: His Majesty's Stationery Office. 6d.

Industrial Health Research Board. *The effect of eyestrain on the output of linkers in the hosiery industry.* By H. C. Weston & S. Adams. (Industrial Health Research Board Report No. 40.) 1927, London: His Majesty's Stationery Office. Illustrated.
An industrial survey.

Industrial Health Research Board. *On the relief of eyestrain among persons performing very fine work.* By H. C. Weston & S. Adams. (Industrial Health Research Board Report No. 49.) 1928, London: His Majesty's Stationery Office. Illustrated. Bibliography.
An industrial survey.

Industrial Health Research Board. *Further experiments on the use of special spectacles in very fine processes.* By H. C. Weston & S. Adams. (Industrial Health Research Board Report No. 57.) 1929, London: His Majesty's Stationery Office. Illustrated. Bibliography.
An industrial investigation.

Industrial Health Research Board. *Manual dexterity: effects of training. I. Transfer of training in manual dexterity and visual discrimination.* By E. M. Henshaw, P. Holman & J. N. Langdon. (Industrial Health Research Board Report No. 67.) 1933, London: His Majesty's Stationery Office.

Industrial Health Research Board. *The effects of conditions of artificial lighting on the performance of worsted weavers.* By H. C. Weston. (Industrial Health Research Board Report No. 81.) 1938, London: His Majesty's Stationery Office.
An industrial investigation.

Industrial Health Research Board. *The relation between illumination and visual efficiency—the effect of brightness contrast.* By H. C. Weston. (Industrial Health Research Board Report No. 87.) 1945, London: His Majesty's Stationery Office. 9d.
Report of an original laboratory investigation.

Industrial Health Research Board and the Illumination Research Committee of the Department of Scientific and Industrial Research. Joint reports. *The relation between illumination and efficiency in fine work (typesetting by hand).* By H. C. Weston & A. K. Taylor. 1926, London: His Majesty's Stationery Office.
An industrial investigation.

Industrial Health Research Board and the Illumination Research Committee of the Department of Scientific and Industrial Research. Joint reports. *The effect of different systems of lighting on output and accuracy in fine work (typesetting by hand).* By H. C. Weston & A. K. Taylor. 1928, London: His Majesty's Stationery Office. Illustrated.
An industrial investigation.

Industrial Health Research Board and the Illumination Research Committee of the Department of Scientific and Industrial Research. Joint reports. *The effect of lighting on efficiency in rough work (tile pressing).* By S. Adams. 1935, London: His Majesty's Stationery Office.
An industrial investigation.

Industrial Health Research Board and the Illumination Research Committee of the Department of Scientific and Industrial Research. Joint reports. *The relation between illumination and industrial efficiency. I. The effect and size of work.* By H. C. Weston. 1935, London: His Majesty's Stationery Office.
Report of a laboratory investigation.

Medical Research Council. *First report of the Miners' Nystagmus Committee.* (Medical Research Council Special Report Series No. 65.) 1922, London: His Majesty's Stationery Office.

Medical Research Council. *Second report of the Miners' Nystagmus Committee.* (Medical Research Council Special Report Series No. 80.) 1923, London: His Majesty's Stationery Office.

Medical Research Council. *Third report of the Miners' Nystagmus Committee.* (Medical Research Council Special Report Series No. 176.) 1932, London: His Majesty's Stationery Office. 9d.

Medical Research Council. *Researches on the measurement of human performance.* By N. H. Mackworth. (Medical Research Council Special Report Series No. 268.) 1950, London: His Majesty's Stationery Office. 4s. Illustrated Bibliography.

Weston, H. C. *Sight, light and efficiency.* 1949, London: Lewis. 42s. Illustrated.
A discussion of the importance of illumination in the visual problems of industry.

Wright, W. D. *Photometry and the eye.* 1949, London: Hatton Press. 7s. 6d. Illustrated.
A discussion of the importance of the properties of the human eye in photometry.

4. Development of the Eye

Mann, I. *The development of the human eye.* [2nd ed., 1949.] Reprinted 1951, London: British Medical Association. 45s. Illustrated. Bibliography.
The standard work on the subject.

Mann, I. *Developmental abnormalities of the eye.* 1937, Cambridge: University Press. 50s. Illustrated. Bibliography.

5. General Physiology

Lovatt Evans, C. E. *Starling's principles of human physiology.* 11th ed., 1952, London: Churchill. 52s. Illustrated. Bibliography.
Book IV. Chapter XXV, by H. Hartridge, deals with vision.

Winton, F. R. & Bayliss, L. E. *Human physiology.* 3rd ed., 1948, London: Churchill. 25s. Illustrated.
Chapter XIV, part I, by R. J. Lythgoe, revised by K. Tansley, gives the best available short account of the physiology of vision for medical students.

6. Physiology and Psychology of Sensation

Adrian, E. D. *The basis of sensation: the action of the sense organs.* 1928, London: Christophers. 7s. 6d. Illustrated.
A short account of the physiology of the sensory end-organs and especially of the sensory nerves.

Adrian, E. D. *The physical background of perception.* 1947, Oxford: Clarendon Press. 12s. 6d. Illustrated.
Waynflete Lectures for 1946. The second lecture deals with the sense organs including the eye, and the fourth with sight and hearing.

Parsons, J. H. *Introduction to the theory of perception.* 1927, Cambridge: University Press. 18s. Illustrated. Bibliography.
One of the classical works on sensation and perception throughout the animal kingdom.

Rawdon-Smith, A. F. *Theories of sensation.* 1938, Cambridge: University Press. 10s. 6d. Illustrated. Bibliography.
Section I is devoted to vision.

Sherrington, C. S. *The integrative action of the nervous system.* 2nd ed., 1947, Cambridge: University Press. 25s. Illustrated. Bibliography.
Some classical observations on flicker and binocular fusion are to be found in the last chapter.

Vernon, M. D. *A further study of visual perception.* 1952, Cambridge: University Press. 35s. Illustrated. Bibliography.

Continued at foot of page 79

INDIVIDUAL DIFFERENCES IN COLOUR VISION

An Essay Review*

ROBERT C. GRAY

Department of Applied Physics
University of Glasgow

Boyle was the first writer to mention colour-blindness; he wrote (1689) of a young gentlewoman who described a colour as " dark or dirty " when to Boyle the colour was " a fine and glossie red "; she obviously had darkened-red vision and would now be called a protanope. In 1777 Huddart described the vision of a man to whom green and red appeared the same as yellow; he was a deuteranope; he had two brothers with the same defect, so that it seemed to be hereditary, but his parents had normal vision. In 1778 Whisson gave a three-generation pedigree of a family in which five males and one female could distinguish only two colours, yellow and blue; they were deuteranopes.

During the 19th century the subject was studied extensively, and by the methods of detection then available it was found that about 4% of men and a very much smaller percentage of women had noticeable colour defect. There were obviously at least two types of colour vision defect:

i. *Protanopia*, a partial colour-blindness, in which the individual's eye is insensitive to long-wave red radiations. Such a defect must be associated with disturbance of the relative sensitivity of the eye to other radiations; and in the extreme case green and red of the normal are replaced by the intermediate sensation of " yellow ", which may not necessarily be the same sensation as the yellow of the normal.

ii. *Deuteranopia*, in which the eye is not in any way blind, since it is sensitive to all the radiations of the normal spectrum, but in which the red and green sensations are weakened, or unbalanced, usually with weakened green, or in which the red and green sensations are replaced completely by a yellow sensation.

Thus in protanopia the vision is weak for perception of red and in deuteranopia the vision is usually weak for perception of green. From this arose the trichromatic theory of colour vision, which postulated that the eye had three sets of receptors, each sensitive to radiations giving rise to one of the three sensations, red, green and blue. This theory has now been abandoned, and Dr Pickford's work gives additional evidence that this theory cannot possibly fit the facts.

The laws of inheritance of colour-blindness were known in the 19th century, but they were not explicable until the discovery, early in the 20th century, of the sex chromosomes, of which the male cell contains one X-chromosome and the female cell two X-chromosomes. It was now evident that the gene for colour vision (or colour-blindness), like the gene for haemophilia, was carried by the X-chromosome, and an explanation of the low incidence of colour-blindness in women, compared with that in men, was available.

The development of pseudoisochromatic cards for colour-vision testing, especially the Ishihara cards, made it possible to detect a number of people whose green-red vision was weak, especially under conditions of low illumination, but who were generally unaware of the defect; the total percentage of colour-blind men detectable had now increased to 8% (Waaler and von Planta), and that of women to 0.44%. Four separate defects were now recognized: the P-group with shortened red (i.e., completely blind to the radiations which are deep red to the normal eye, of wavelengths greater than 620 mμ.), consisting of (i) protanopes (1% men), with no red-green vision distinct from yellow, (ii) protanomalous (1% men), with some red-green vision; and the D-group, with no shortening of the red, consisting of (iii) deuteranopes (1% men), with no red-green vision distinct from yellow, and (iv) deuteranomalous (5% men), with some red-green vision. Dr Pickford's work shows that these groups may be further subdivided, and that there are large variations among the normals.

These four defects have been found to be inherited true. Further, by study of the vision of colour-blind women, all of whom must have a double defect, it has been shown that deuteranomaly is dominant to deuteranopy, and that protanomaly is dominant to protanopy. Also, women who had the gene for deuteranopy or deuteranomaly in one X-chromosome, and the gene for protanopy or protanomaly in the other X-chromosome, were found to have almost normal colour vision. Since there are 6% of men in the D-group and 2% of men in the P-group, there should be 0.36% of women in the DD-group and 0.04% in the PP-group; thus the number of colour-blind women should be 0.40% of the population, a figure that agrees well with 0.44% obtained from Waaler's and von Planta's surveys.

---

* **Individual Differences in Colour Vision**
R. W. Pickford. London: Routledge & Kegan Paul Ltd., 1951. xviii + 386 pages; 1 plate; 33 diagrams. 22 × 14 cm. £1 10s.

PHYSIOLOGY OF VISION BIBLIOGRAPHY

Continued from page 78

Vernon, M. D. *Visual perception.* 1937, Cambridge: University Press. Illustrated. Bibliography.

7. Nutrition

Medical Research Council. *Vitamin A requirements of human adults.* By E. M. Hume & H. A. Krebs. (Medical Research Council Special Report Series No. 264.) 1949, London: His Majesty's Stationery Office. 3s. Illustrated. Bibliography.
Account of an experiment on vitamin A deprivation in man with much evidence as to its effects on dark-adaptation.

8. Genetics

Grüneberg, H. *The genetics of the mouse.* 1943, London: Cambridge University Press. 2nd ed., 1952, The Hague: Nijhoff. 90s. Illustrated. Bibliography.
Section on developmental and genetic diseases of the eye.

9. General Biology

Tinbergen, N. *The study of instinct.* 1951, Oxford: Clarendon Press. 25s. Illustrated. Bibliography.
Incidental facts about vision and the effects of visual stimuli in many different species.

Dr Pickford's approach to the subject of colour vision is that of the experimental psychologist. He begins with an introductory chapter on the problems of colour vision, and details are given of the several theories that have been from time to time suggested in order to explain colour vision and its many variations. In succeeding chapters are described the various pieces of apparatus developed, all of them interrelated but giving independent results; the results obtained are presented in an orderly way, and the conclusions that may be drawn are stated clearly.

The first test was a Rayleigh equation experiment, in the first part of which two rotating disks were used, one yellow-black-white, the other red-green, and the subject matched his sensations both for illumination and hue. In the second part, a yellow-blue disk was matched against a black-white disk. It was noticed that the matching point gave only a small portion of the information desired, and the experiment was extended to measure the range of settings that the subject would accept as a match. Two figures for each subject were now available: his deviation, i.e., the amount by which the mid-point of his matching range differed from the mean mid-point of the normals, and the extent of his range of matches. These give Dr Pickford a means of dividing his subjects into (i) normal, with small range and small deviation, (ii) colour-weak, with moderately large range and small deviation, (iii) deviants, with small range, but moderately large deviation, (iv) anomalous, with small range, but very large deviation, and (v) colour-blind, with very large range, and small, moderate or even large deviation.

Further division of these classes is made by the author. The colour-weak may be weak in one, two or more colours; the deviants may show their variations in one colour, or in two colours together such as red and blue, but not in complementaries; the anomalous may have green-anomalous colour vision, red-anomalous without darkened red, or red-anomalous with darkened red; the colour-blind may be divided into moderates and extremes, and again into protanopes (with darkened red), deuteranopes (with red of normal brightness), and deviant deuteranopes.

The red-green tests support Houstoun's conclusion that the ability to distinguish red from green is subject to variation among the ordinary population, so that the main group includes the normals, colour-weak and deviants, and the red-green blind, including the anomalous, tend to form a well-marked separate group connected to the main group by a very small bridge. The Ishihara test fails the colour-blind and the anomalous, and if applied strictly may fail a proportion of the colour-weak. The heterozygous woman, or conductor, who has one colour-normal X-chromosome and one colour-defective X-chromosome, tends to be red-green weak. Green is more frequently weak than red.

No anomalous or colour-blind subjects were found in the yellow-blue tests. Yellow and blue are found weak more frequently than red or green. Blue weakness is always accompanied by a tendency, when reference is being made to "near shades" of blue and green, to call them green rather than blue, and also by a tendency to say that a yellow object is white. In yellow-blue vision, women do not differ significantly from men, so that yellow-blue weakness cannot be sex-linked. Weakness in yellow-blue vision is not necessarily accompanied by weakness in red-green vision; nor is red-green weakness accompanied by greater yellow-blue sensitivity.

The subsequent tests were designed to confirm these results in greater detail, and to meet certain objections that had been made to the arrangement; several additional points of value were recorded. The apparatus using rotating disks was replaced by an anomaloscope, in which the colour of a spot of light could be changed continuously from red to green, or from yellow to blue; the results obtained confirmed the previous findings. A specially designed brightness test showed more definitely that there was no correlation between colour discrimination and brightness judgement. The objection that polychromatic filters had been used was met by a test in which Ilford monochromatic filters were employed; no significant differences in the results were found. To meet the objection that the subjects were being tested in only four colours—red, green, yellow and blue—the experiment was extended to the use of additional filter-pairs with colours of violet and complementary yellow-green, orange and complementary blue-green, and green and complementary magenta; the results showed that defects in sensitivity in the intermediate colours are due to weaknesses in one or more of the four colours, red, yellow, green and blue. Darkening of the violet end of the spectrum was found in three men and three women. In this experiment the subject was classified as "dark" or "fair", and no relation between colour vision sensitivity and skin or hair pigmentation was found.

Further tests were carried out in which the subjects matched, in succession, a red and yellow mixture with an orange standard, a yellow and green mixture with a yellow-green standard, a green and blue mixture with a green-blue standard, and a blue and red mixture with a violet standard. It was found that the distinction between blue and green is the most variable, among normals, of the four distinctions tested, that red-blue distinction is next in variability, followed by red-yellow and yellow-green. Tests with binocular vision showed that for normal subjects yellow and blue are difficult to combine, and red and green more so, when one colour is presented to each eye; this supports a four-colour theory of the Hering or Houstoun type in which opposite colours would be expected to combine by a neutralizing process in the retina of one eye, but not when presented separately to the two eyes.

Dr Pickford's descriptions of the attitudes of his many subjects to the tests, and their responses to requests for names of colours presented, are very interesting and valuable, and at times humorous. He gives much valuable advice on the attitude of the tester: it is peculiar to colour vision tests that ". . . the subject may learn to do the tests by guides and cues external to the capacities being tested"; and the tester has to be more of a clinician than a tester to penetrate the numerous mechanisms utilized by the colour-blind subject to protect himself against the suspicion that he is not perfectly normal.

The future worker on this subject will find in this book abundant help to enable him quickly to obtain trustworthy results, and to avoid the many pitfalls, physical and psychological, that beset the researcher in this field.

Book Reviews

Colour in Theory and Practice

Edited by *H. D. Murray*. New edition, revised, enlarged and rewritten. London: Chapman & Hall Ltd., 1952. xiii + 360 pages; plates; figures. 25 × 19 cm. £3 10s.

"Colour" is sometimes used in a physical sense to denote a property which "coloured" objects have in common: that of altering the spectral composition of the light by which they are seen. This meaning is implicit in the layman's insistence that grass is always "green"—even at night in the light of a sodium vapour street lamp. The colour name of a substance represents an attempt to describe in everyday language a permanent property; a property which can be described adequately only by an absorption spectrum. An understanding of these matters is helped by a study of the relations between the absorption of light and the structure of matter, subjects which are the concern of the first section of this book, "Physical and chemical aspects".

The colour names are also used—and to the purist this is their only legitimate use—to describe sensations evoked by stimulation of the visual apparatus, whether by light or otherwise. "Colour" in this sense forms the subject of the second section, "Physiological and psychophysical aspects".

Yet a third meaning for "colour" is apparent in the description of light itself as coloured. This aspect of meaning which is ultimately expressible in terms of emission spectra is treated in the third section, "Light sources and colorimetry".

These three connotations of "colour" form the broad framework of this book. A brief final section on miscellaneous aspects is of a slight nature and adds little to the value of the whole.

The first section is not well co-ordinated with the others. Thus in the fragmentary chapter on photochemical effects and in the otherwise adequate chapter on absorption of light, no mention is made of the visual pigments. These could have provided a more apposite medium for exposition, leading the reader along a smoother path to the subsequent chapters on vision. The difficult subjects of electronic wave forms and colour and chemical constitution, however, are handled in an assured and stimulating manner. They help to redeem a section written in a disjointed and somewhat perfunctory style.

The next section comprises a lucid account of the visual apparatus and its functioning in normal and defective colour vision. Nicely interwoven through this account is an exposition of the specification and measurement of colour. The line illustrations are particularly instructive and are worth careful study. It is unfortunate that two figures referred to in the chapter entitled "An explanation of the visual processes" are missing. Two of the colour plates in this section may also be criticized: that illustrating simultaneous contrast is not convincing, while the other, illustrating the three attributes of colour sensation, does not clearly distinguish between luminosity and saturation differences.

The third section is full of useful information on light sources and colorimetry, and its value to practical workers is enhanced by a number of appendices.

The book, which opens flat, is excellently produced, the type is very pleasant and the illustration lavish. The line drawings and half-tone plates are of very high standard; the same cannot be said of the colour illustrations. Whereas those illustrating additive and subtractive colour mixture and after-images are excellent, two others, already mentioned, fail in their object. The remaining four need not have been in colour at all. Their replacement by line drawings might have reduced the cost of this expensive production.

The book is good; good enough to make one wish that it had been a little better. Though the individual sections are authoritatively and lucidly presented the work as a whole lacks unity. In producing it the services of no fewer than seven experts were enlisted by the editor. Though he cannot be said to have succeeded in his difficult task of integration the volume is nevertheless one whose quality has demanded assessment against exacting standards.

H. J. A. Dartnall

Clinical Pathology of the Eye: a Practical Treatise of Histopathology

Bernard Samuels & Adalbert Fuchs. London: Cassell & Co. Ltd., 1952. xviii + 420 pages; 44 plates; 226 figures. 28 × 19 cm. £7 7s.

This volume is a considerable addition to ophthalmological literature and in a sense is original in its presentation. As its name implies it is not merely a description of post-mortem appearances but a close correlation of these with the appropriate clinical pictures. More than any other organ the eye lends itself to such eminently logical treatment, for with the transparency of its media and its ready access to minute and exact examination with the many optical devices now at the disposal of the clinician, most of the structural manifestations of disease can be observed with considerable accuracy during life. This makes the correlation of the clinical pictures with the histological appearance relatively easy, for the one leads to the other without any hiatus and the combination of the two studies becomes one logical whole. This clinico-pathological combination is seen to its best advantage in the discussion of retinal diseases, for in this chapter the many appearances seen ophthalmoscopically—white, yellow or red patches and dots, streaks and strands which so often puzzle the student and frequently mislead the clinician—are discussed on the basis of the microscopic findings apart from the diseases wherein they usually occur. It is a useful and rational approach not lacking in originality.

On this basis the volume covers the entire pathology of the eye. The diseases of the cornea, sclera, uvea, lens, retina, vitreous and optic nerve are treated systematically, and there are chapters on myopia, tuberculosis, syphilis, postoperative findings, injuries and tumours. The diseases of the conjunctiva are not included nor are those of the orbit, lids and ocular adnexa; this is perhaps unfortunate since the space allotted to them in texts of general pathology is meagre and their importance from the point of view both of the ophthalmological student and practitioner is considerable.

The pathological teaching throughout the book is based on the teaching of Ernst Fuchs of Vienna, the greatest exponent of ocular pathology in the latter part of the last, and the beginning of the present century. Both authors studied the subject in his laboratory, and the present work is a combination of the lectures they have since delivered, the one at the New York Eye and Ear Infirmary and the other at the Allgemeines Krankenhaus in Vienna; with both, however, having derived from the same source and depending on the same inspiration, the combination has been easily made. Perhaps for this reason the general impression given is of a past era rather than of present advances. The classical view-point is everywhere apparent. To take one example, the elongation of the globe in myopia is stated to be the result of diminished resistance of the sclera to the normal intra-ocular pressure and the choroidal changes are said to be due to stretching. There is no suggestion of the more modern view that the shape of the eye and the pathological changes therein may more often be genetically determined and are not of purely mechanistic origin.

The same philosophy is extended to the illustrations, which are abundant. With one exception these are drawings, some of them, it is true, of considerable merit, although not all of them are original to the present book. The authors suggest that drawings are more valuable to the student than photomicrographs since details which are of importance can thus be emphasized. In a sense this may be true; but in many cases the illustrations depart somewhat from the appearance of the histological slide and the exaggeration of specific details may not be an unmixed blessing to the student if he studies pathology as he ought to—with one eye on the microscope and the other on his reference book. The modern pictures of histological specimens need lack nothing in clarity.

On the whole the book is a most interesting presentation of its subject and will be read with interest and profit, not least because of the simplicity and directness of the presentation and the clarity of the writing.

Stewart Duke-Elder

Tests for Colour-Blindness

Shinobu Ishihara. 10th ed. Tokyo: Nippon Isho Shuppan Co. Ltd.; London: H. K. Lewis & Co. Ltd., 1951. 27 pages; 38 plates. 21 × 15 cm. £3 15s.

The issue of a 10th edition of this famous test for colour-blindness calls for a statement of its capacities and limitations so that intending users of the test may know what to expect of it. The 9th edition was produced in Britain, and was by many considered slightly inferior to the 8th Japanese edition; L. H. Hardy and his co-workers found no marked differences between the 5th, 7th and 9th editions. The present or 10th edition is again printed in Japan, but it has not been available long enough for any adequate empirical studies of its validity and reliability to have been made, or for comparisons with the earlier editions to be published. Professor Ishihara's rather arbitrary classification of colour vision defects remains in the notes and instructions for this edition, but is out of date and needs no comment.

Six new plates have been added, numbered 26, 27, 34, 35, 36 and 37. Plate 1 is a " joker " which can be read by all but the totally colour-blind, and the reviewer has met only one subject who could not read it among about 1,200 tested with this and other tests. Plate 38 is a similar " joker ", for non-readers, using " wavy lines " instead of numerals, and these two plates may be used for showing the subject what to do in the test. Plates 2–5 have pink figures on greenish backgrounds, while plates 6–9 have green figures on pinkish backgrounds. All eight of these plates have alternative readings for normal and colour-blind subjects, and those with single-figure numbers as shown in plates 2, 3, 6 and 7, are more reliable, because less confusing than the others. Many people have thought that those plates which give most mis-readings and create most difficulties for the bulk of the population, such as plate 9, must be the most sensitive and best, but careful researches have shown that the mis-readings are simply due to the confusing nature of the figures to be read.

Plates 10–13 have pink figures on greenish backgrounds, and plates 14–17 have green figures on pinkish backgrounds, but without alternative readings for the colour-blind. These are more reliable than the plates with alternatives, because they are less confusing and again those with single-figure numbers to be read are better than those with two-figure numbers.

None of the plates 2–17 can give any reliable indication of the type of the subject's defect, and the degree of defect is not revealed satisfactorily by the number of errors made, either for the normal or for the colour-blind subject, but the colour-blind make on the average about five times as many errors as the normal on these 16 plates.

Plates 18–21 have figures camouflaged for the normal but legible to many of the defectives. It is encouraging to the colour-blind subject to be able to read these plates, when the normal cannot, and also to be able to see different figures from those read by the normal on plates 2–17, but it is of little or no help to the tester.

Plates 22–25 show figures in two colours on dark grey backgrounds. Protanopes are supposed to be able to read the purple figures, whereas for them the red figures are meant to be indistinguishable from the backgrounds. Deuteranopes are supposed to be able to read the red but not the purple figures. Broadly, the parallel distinction holds for red- and green-anomalous subjects, but the red-anomalous with red vision of normal brightness are classified as deuteranopes " green-blind " if they can read the red figure better than the purple. These plates were very unreliable and of low validity in the 8th and 9th editions. A short series of trials suggests that in the 10th edition these four plates are very weak for distinguishing protanopes because the backgrounds have been made too dark and protanopes can read the red figures quite easily.

At this point we should turn to the end of the test and start with plate 38, working backwards to plate 26. Then it is found that the essential principle of construction already explained for the first 25 plates is duplicated in a shortened form with " wavy-line " patterns for non-readers, with whom the tester would proceed from plate 38 to plate 26 and omit the first 25. The " wavy-line " patterns are quite good for this purpose, but have no advantage over the other plates for any subject who can read. In the 10th edition some interesting additions have been made to the plates for non-readers, which give alternative readings. The alternatives for plates 34–37 are normal or defective " readings ", while plates 26 and 27 are supposed to distinguish protanopes and deuteranopes, but are not reliable.

The Ishihara test was a masterpiece when first produced, and has been widely used. If one realizes its limitations it still is a valuable test, but for efficient and reliable work on individual differences in colour vision it must be supplemented by tests of the anomaloscope type, which are better than any with pseudoisochromatic plates. The limitation, however, proved by Chapanis, Hardy and others, is that the Ishihara test discriminates successfully between the normal and the major red-green defectives as a group, without making other reliable distinctions. Minor variations among the normal are not indicated efficiently by the test, and no reliance can be put upon it to differentiate types or degrees of colour-blindness or of anomalous colour vision.

When the 8th edition is used for testing, the average number of errors made with plates 2–24 inclusive falls off steeply for normal subjects, from 0 or 1 to about 8 or 10, and the average number for major defectives rises steeply from 14 or 16 up to 24 errors. The failure level was therefore at about 12 errors on these plates. For the 10th edition the level of failure is probably about the same on the 24 plates excluding plates 1 and 26–38. This criterion is on the safe side and may reject a few colour-weak and deviant normal subjects, but where life depends on colour vision it is wise to be strict.

A short preliminary study of the 10th edition, with defectives whose colour-blindness is well known to the reviewer, suggests that this edition is less reliable than the 8th, and that the 9th edition falls between them. The additional plates are interesting but do not add much of value, and it is still true that a short test of six plates and a " joker ", composed of plates 1, 10, 11, 14, 15, 18 and 23, would be quicker, more reliable and of greater validity than the whole test.

R. W. Pickford

Color in Business, Science, and Industry

Deane B. Judd. New York: John Wiley & Sons, Inc.; London: Chapman & Hall Ltd., 1952. ix + 401 pages; 106 figures. 23 × 15 cm. £2 12s.

This is a tantalizing book. The style rises from such degradations as ". . . it is supposed that there are red-reporting cones that ought to have long-wave sensitive stuff in them; there are green-reporting cones that ought to have middle-wave stuff in them; and violet-reporting cones that ought to have short-wave stuff " to nearly epigrammatic levels: " New tools are constantly being devised, and old ones sharpened ". One reads about " spectroradiometers " and vacillates between verbal associations with the Aurora borealis and Rorschach tests. The reader is presented with a lucid diagram illustrating additive colour mixture (fig. 15) but also with a poor colour reproduction of a " Munsell red " (fig. 93). Is " psychophysics " the physics of the soul? Why is " perception " used as a synonym for " sensation " ?

While disturbed by some of the errors (chicken retinas *do* contain visual purple, rods *are* found in human retinas within 1° of the fixation area, and there are cases of total colour-blindness or cone monochromatism which the literature has *not* stamped as acquired), the reader is lost in admiration for Dr Judd's profound knowledge of all the technicalities of colorimetry. The explanations of techniques and apparatus are masterly, and given with an air characterized by the absence of all pedantry. Due credit is given to British workers. Le Grand's recent contribution to the evaluation of trichromatic data is missed, however, because, like Dr Judd, he is searching for clues in the realm of adaptation. The volume is invaluable to all who are interested in the physical aspects of colour: useful data on pigments, tristimulus values for the spectra of standard sources B and C, and an extensive and up-to-date bibliography render this book unique in its self-sufficiency.

R. A. Weale

Vision and the Eye

M. H. Pirenne. London: Chapman & Hall Ltd., 1948. (Frontiers of Science Series.) xx + 187 pages; 4 plates; 99 figures. 19 × 14 cm. 12s. 6d.

As Dr Stiles remarks in the foreword, this little book provides a valuable introduction for those embarking on research in vision. Its arguments are well presented so that they are easy to follow: the book might well have been named "Vision without tears". A wide range of subjects is discussed, adaptation forming the one exception. This is somewhat surprising since it can be described quantitatively and thus falls within the author's avowed physiological approach. One is grateful for Dr Pirenne's total abstinence as far as the orgies of colour vision arguments are concerned, and moved by his tactful loyalty shown when he rejects Hecht's theory of visual acuity.

There are—inevitably—points on which opinions differ; for example, should the term "quantum effectiveness" be used (p. 53) before it is explained, and Stiles's response curves be called "broad" (p. 158)? The latter appear to be so only when plotted on logarithmic ordinates. These are merely minor points. In the reviewer's opinion, however, a dilemma of some importance may arise in the careful reader's mind concerning Dr Pirenne's main contribution to vision, as given in this book. There is no doubt that the act of vision, being initiated by a photochemical reaction, is quantized. On the basis of some old and only approximate results, Chalmers suggested as far back as 1916 that the absolute visual threshold corresponded to the energy of one quantum. Using the method of frequency-of-seeing and Poisson's distribution, Dr Pirenne arrives at a result of 5–7 quanta, which is said to agree with the absolute threshold measurement of 5–14, when the requisite corrections have been applied to the raw data. This agreement is at the root of the dilemma. One of the corrections involves the assumption that the retinal visual purple absorbs as much as 20% of the incident light (p. 82). Yet the scotopic sensitivity data shown on p. 52 are in manifest disagreement with the energy absorption spectrum corresponding to this concentration, but agree with an absorption of only 5%. The "upper limit" of 5–14 is thus reduced to 1–3, a value in good agreement with the results of other measurements. The reviewer believes that this discrepancy between the frequency-of-seeing and other data casts doubt on the applicability of the mathematics underlying the former. Crozier has recently shown that the foveal frequency-of-seeing function cannot obey a Poisson distribution; tests for the periphery are not extant. Moreover, Lee, Finch & Pounds have provided data which show a regular time variation of the absolute threshold, a matter which has not been considered in *Vision and the eye*. It is only reasonable to expect that these points may be elucidated in future editions, which, one does not hesitate to say, Dr Pirenne's well-illustrated book certainly deserves.

R. A. Weale

A Further Study of Visual Perception

M. D. Vernon. Cambridge: University Press, 1952. xi + 289 pages; 30 figures. 22 × 14 cm. £1 15s.

Many of those who interest themselves in the physiology of vision regard the activities of the experimental psychologists with mild impatience—if indeed they are aware that such work even exists. That there is some justification for this attitude is shown by the opening of Miss Vernon's introductory chapter, in which she explains that her previous book on this subject, published in 1937 and largely concerned with the work of the Gestalt psychologists, no longer gives a satisfactory picture of the psychology of visual perception. She suggests that the Gestalt theory no longer provides "a fundamental basis for the understanding of the nature of perception, but rather a one-sided exaggeration of certain features by no means the most important in perceiving as we ordinarily experience it." A few pages further on she makes another remark about "that curious emphasis on the unimportant which has so often been the bane of psychology." Anyone who is not an experimental psychologist but who has had to read and try to assess the importance of even a small part of their enormous literature cannot but greet such statements in such a context with unkind delight.

Nevertheless, workers on the physiology of vision must sooner or later consider sensation whether they will or no, and much has been discovered by experimental psychologists of which vision physiologists are woefully ignorant and of which they ought to know. Miss Vernon's admirable book could, therefore, help to bridge a serious gap if the physiologists of vision would read it with an open mind, for she reviews the results obtained in psychological laboratories over the last fifty years in a critical and helpful manner.

The argument of this book is based on the astonishing way in which a final percept can differ from that which would be expected from a consideration of the physical stimuli which produce it. Thus, for instance, no one who sees a penny lying on a table doubts that he is looking at a circular object, although the actual retinal image is nearly always an ellipse. Miss Vernon believes that such interpretations of sets of stimuli are made because of the individual's need to maintain the maximum stability, endurance and consistency possible in his apprehension of the external world. Such stability must, however, be combined with enough perception of important external changes to ensure suitable reactions on the individual's part when necessary. In developing her argument she considers the general nature and development of perception and the recognition of form, space and movement. There is also a chapter on the fascinating problems of size, shape and colour constancy, and another on the factors governing individual variations in visual performance.

While it is true that the vision physiologist needs some knowledge of recent psychological results, one also feels that it would be advantageous if the experimental psychologist were more familiar with modern work in sensory physiology. With the exception of Hartridge's by no means reliable *Recent advances in the physiology of vision*, the latest (and only) physiological work quoted by Miss Vernon is Adrian's *Basis of sensation*, published in 1928. Results such as those of Granit in his work on the retina or of the American workers on the visual centres of the brain suggest that the picture supplied by the sense organs is by no means as simple as is implied here. It is probable that some, though by no means all, of the variability of sensation discussed arises in the periphery and need not be explained on a psychological basis. This is, however, a relatively unimportant failing in an excellent and well-balanced book which provides a most helpful account of the current problems of experimental psychology and of the way in which they are being attacked.

Katharine Tansley

Principles of Human Physiology

Originally written by *E. H. Starling*. 11th ed. by *Charles Lovatt Evans*, with chapters on the special senses by *H. Hartridge*. London: J. & A. Churchill Ltd., 1952. xii + 1210 pages; 709 figures. 25 × 16 cm. £2 12s. 6d.

The British Encyclopaedia of Medical Practice. Volume 10. Pituitary Gland Diseases to Retina Diseases

Edited by *Horder*. 2nd ed. London: Butterworth & Co. Ltd., 1952. xv + 718 pages; 5 plates; 117 figures. 25 × 17 cm. £3 5s.

Although it is not possible to review these books fully, they each contain a chapter with a bearing on vision. Reviews have therefore been included of these chapters: in *Principles of human physiology*, chapter XXV, "Vision", by H. Hartridge; in *The British encyclopaedia of medical practice*, "Refraction, practical methods", by J. H. Doggart.

Few men have written more on vision than Professor Hartridge and, in this latest edition of what has become his part of "Starling", he has more than consolidated the position which he occupies in the subject. When such a ramified section of physiology is surveyed, people may differ in the amount of emphasis they would like to give to various parts of it. For example, the impression given by this chapter is that work on stereoscopic vision does not seem to have progressed since the earlier editions of "Starling"; in fact, the various "alley" experiments have led to distinctions between stereoscopic and physical space, the mathematics of which might repel many, but fascinate a few students. Again, the vastness of the field makes it inevitable that errors have crept in, as the assiduous student will discover on consulting original papers: thus, Dartnall has never claimed to have discovered six visual

pigments in the tench retina, as described on pages 446–7. Moreover, in one of his own publications cited in this chapter, Professor Hartridge has described the discovery of a fixation point for light of wavelength 5,200 A., the present omission of which may perhaps be noted by those familiar with his work. These, however, are matters of small importance. The section is written with the verve and the capability for logical presentation of relevant matter which one has come to expect from this writer. This chapter of *Principles of human physiology* can be recommended particularly to the medical student interested in the history of vision, because no other contemporary writer would take the trouble to show him references to, and graphs from, long-forgotten papers.

The well-illustrated article on refraction, which is included in the 10th volume of *The British encyclopaedia of medical practice*, adequately covers most of the tests used in refraction, without entering into their theoretical basis. A description of preliminary investigations—which can be successful only after sufficient experience on the practitioner's part—is followed by a survey of visual acuity tests: the fact that Landolt's C is no longer mentioned testifies to the high standard of the contribution. A description is given of the methods and advantages of retinoscopy; emmetropia, hypermetropia, myopia and astigmatism are dealt with in some detail. The description of the duochrome test suffers from the fact that the criterion of acuity (sharpness of outline) is replaced by one of blackness: a protanope may well see the letters on the green background as darker than those on the red, even when the latter are " in focus ", since his reduced sensitivity to red light gives rise to a lower contrast. The chapter concludes with brief references to subjective tests for near vision and to special devices such as prisms, the streak retinoscope and other refractometers. It is a pity that this excellent review relegates special apparatus to a back seat: the omission of any reference to Fincham's coincidence optometer is regretted, since the objective measurement of the patient's refraction to an accuracy of 0.01 D. is both easy and elegant.

R. A. Weale

History of the Royal Medical Society, 1737–1937

James Gray, edited by *Douglas Guthrie*. Edinburgh : University Press, 1952. xi + 355 pages ; illustrations. 25 × 16 cm. £2 2s.

This is a book replete with interest from the first page to the last in that it is a detailed record of the Royal Medical Society of Edinburgh and of its many famous alumni since its inception in 1737 down to the present day.

The author, James Gray, though not himself a medical man, was intimately connected with the Society and its activities for many years, and obtained the material for this history from the minutes themselves, the dissertations communicated and other sources. He died untimely in 1942 but had arranged for the publication of his manuscript. Others have brought the work up to date. In an appendix are the names of the annual presidents for each year from 1764 to 1952. If any confirmation of its accuracy were needed this would be found in the editorship of the renowned medical historian, Dr Douglas Guthrie. Biographical notes are given of more than 600 who have held official posts in, or have been associated with, the Society.

The author describes how the torch of learning came to be handed on from Leyden to Edinburgh in the 17th century ; how Robert Sibbald, a Leyden student, was appointed to the Professorship of Medicine in the Town's College (later, Edinburgh University) in 1685, and how, reciprocally, Archibald Pitcairne of Edinburgh became, in 1692, Professor of Medicine at Leyden. The Edinburgh University Faculty of Medicine was not founded until early in the 18th century (1726), with Andrew Sinclair and John Rutherford (the maternal grandfather of Sir Walter Scott) as " professors of the theory and practice of medicine ".

The Medical Society was founded in 1737. Medical students had met informally for the preceding three years to discuss medical subjects. The first paper read, says Mr Gray, was on canine rabies, on the 20 December 1734. In 1778 the Society was incorporated by Royal Charter.

The author introduces to the reader many famous medical men, some of whom later gained world-wide reputations. Among these are: James Sutherland, the first Professor of Botany at the University, appointed at the princely stipend of £20 per annum, and Robert Eliot, appointed in 1705 as Professor of Anatomy at £15 per annum on condition that he took charge of the rarities in the College and made " ane exact inventar " of them. A later successor was Monro of resurrectionist fame. Several who had been shining lights of the Edinburgh Medical Society came south to London and enhanced their reputation by work in the metropolis. To mention a few: Alexander Russell, after visiting London in 1735, travelled in Turkey and learnt much about plague. In 1755 he became physician to St Thomas's Hospital, and died of typhus in 1768. Mark Akenside went from Edinburgh to Leyden where he took the M.D., and in 1747 settled in London, obtained the M.D. of Cambridge, became a Fellow of the Royal College of Physicians and in 1759 physician to St Thomas's; he died of typhus in the same year as his colleague, Dr Alexander Russell. Gilbert Blane, who served with Admiral Rodney and introduced lemon juice to ward off scurvy among sailors, was in 1793 another of St Thomas's physicians. His students, with customary flippancy, called him " Chilblaine " on account of his aloof demeanour towards them. Charles Murchison was another who came from Edinburgh to St Thomas's and on his work at the London Fever Hospital wrote the famous *Treatise on the continued fevers of Great Britain*.

Astley Paston Cooper studied first at Guy's, then in Edinburgh, and later returned to London and lectured on surgery at St Thomas's. Benjamin Travers, another Royal Medical Society member, was his colleague and together they wrote *Surgical essays*. Among others who came south, the following became physicians to Guy's Hospital: Richard Bright, famous as the describer of what was thereafter known as Bright's disease, and Thomas Addison, of Addison's disease and Addisonian (pernicious) anaemia; also William Saunders, Thomas Skeete, James Curry, and Thomas Hodgkin of Hodgkin's disease (lymphadenoma), who studied with Laennec in 1822 and introduced the use of the stethoscope at Guy's Hospital.

University College Hospital can claim, among the Royal Medical Society's alumni, Robert Liston, who performed the first major operation in Great Britain with ether as a general anaesthetic; also William Benjamin Carpenter, William Sharpey, James Syme and Burdon Sanderson. To St Bartholomew's Hospital came Dyce Duckworth and Thomas Lauder Brunton, who lectured also at the Middlesex Hospital. Sir James Purves-Stewart, author of the standard work on *Diagnosis of nervous diseases*, was physician to Westminster Hospital.

Among notabilities in medical jurisprudence mention must be made of Sir Robert Christison and Harvey Littlejohn. Sims Woodhead, the pathologist, held the Chair at Cambridge and founded the *Journal of Pathology and Bacteriology*. Law is represented by Robert Bannatyne Finlay, who joined the Royal Medical Society in 1861, took the M.D. in 1863, was admitted to the Bar, became a Q.C. in 1882, Solicitor-General in 1895, Attorney-General in 1900 and Lord Chancellor in 1916.

The above is only a sketch, bare and very inadequate, of the many interests of this fascinating volume, which will appeal to Scotsmen scattered over the globe who in their halcyon days met with or were themselves members of the Royal Medical Society.

H. Harold Scott

• • •

Printed in Great Britain by Spottiswoode, Ballantyne & Co. Ltd., London & Colchester

	PAGE
Introduction — *Sir Rudolph Peters*	85
Cytochrome — *D. Keilin & E. C. Slater*	89
Some aspects of the energy transformation in living matter — *H. A. Krebs*	97
Alternative routes of carbohydrate oxidation — *Frank Dickens*	105
Phosphate-transferring enzymes — *M. Dixon & E. C. Webb*	110
Significance of biochemical lesions in the pyruvate oxidase system — *Sir Rudolph Peters*	116
Folic acid and related compounds in the metabolism of micro-organisms — *D. D. Woods*	122
β-Glucuronidase and related enzymes — *G. A. Levvy*	126
The formation of peptides in enzymic reactions — *Charles S. Hanes*	131
Amino-acid decarboxylases — *E. F. Gale*	135
Cholinesterases — *R. H. S. Thompson*	138
Action of drugs on enzyme systems — *J. H. Quastel*	142
Enzymic oxidation of amines — *H. Blaschko*	146
Adaptive enzymes — *Sir Cyril Hinshelwood*	150
Enzymes acting on nucleic acids — *J. N. Davidson*	154
Plasma alkaline phosphatase in disease — *E. J. King*	160

Scientific Editor : *Frank Dickens*

Some Aspects of Enzyme Research

This number of the Bulletin may be regarded as a sequel to the Gowland Hopkins memorial volume, New Currents in Biochemistry.[1] It contains a collection of articles by experts in their fields, many of whom deal with historical perspectives.

The historical approach is one which is out of fashion in much of the world today ; it is felt with some justice that the power of new techniques in biochemistry, such as the use of isotopes, has created advances so great that the older work may safely be neglected ; and indeed that it is more suited to scientific scholars, or perhaps to those who do not wish to be forgotten.

[1] *British Medical Bulletin*, 1948, Vol. 5, No. 4–5.—ED.

Such a view at the least endangers scientific effort, leading as it may to the unconscious and unnecessary repetition of previous work; but, more important, it fails to understand the clear and basic growth of scientific fact and theory without which the true import of contemporary efforts may be judged.

The significance of enzyme study in relation to pharmacology and neurology may be said to date from the work of Dale, Loewi and the Stedmans in regard to cholinesterase; and an article is in this number by R. H. S. Thompson, dealing with the group of enzymes now forming this field, to which he and his colleagues are making major contributions. Nevertheless cholinesterase is an enzyme concerned with nerve transmission rather than with intermediary metabolism. Biochemists, it is true, have believed in the ultimate value of their studies upon fragmented tissue; but even in 1939 few medical men, even few physiologists, accepted the implication of the ideas involved. The wider recognition of the unique importance of *intracellular* enzymes in catalysing specific metabolic paths is of only recent growth and the understanding of their influence on medicine a contemporary study, the pace of which is being forced by the bridges forming between enzyme biochemistry and functional studies, in the broadest sense, to which it is hoped the writer's article on the pyruvate oxidase system can contribute

In fact modern knowledge upon intracellular enzymes does not date back much more than 25 years; it originates in large part from the work of Fletcher and Hopkins in Cambridge in 1906 on lactic acid formation in muscle, of Harden and Young and Neuberg on yeast, together with the rather later studies of Embden, Hill and Meyerhof. Valuable pioneer work on enzymes was also done by von Euler. The early lactic acid work emphasized Hopkins' thesis of the value of following the dynamic changes of small molecular substances; built upon this, Meyerhof's demonstration that glycolysis would proceed in a muscle extract provided the experimental basis for the modern analysis of this complex system of enzymes, and so of the Embden-Meyerhof glycolytic scheme. Extensions of our knowledge of carbohydrate metabolism are reviewed by F. Dickens, who gives an account of his own and other interesting new work on alternative paths, and of the formation of pentose.

In the oxidation field, to D. Keilin should go the credit of understanding how his observations on cytochrome could be crystallized into first ideas of the respiratory chain; this reconciled the pioneer dehydrogenase experiments of Thunberg and the conflicting views of Warburg and Wieland upon tissue respiration; it is fitting that the developments of the cytochrome story should be described by D. Keilin with E. C. Slater.

J. H. Quastel reviews inter alia his early work (with colleagues) on the competitive effect of malonate on the succinic dehydrogenase in bacteria, which was used in the fundamental studies of Szent-Györgyi and of Krebs. The principle of competitive inhibition (Woods and Fildes) has elucidated the action of sulphonamides on bacteria, and has formed an inspiration for much work since. The developments of *p*-aminobenzoic acid in the folic acid field are reviewed by D. D. Woods in this Bulletin. H. A. Krebs puts together the present conception of final disposal of C_2 units in metabolism by the tricarboxylic acid cycle, which he has done so much to formulate, together with recent work by others upon the oxidation of fatty acids, glycerol and the degradation of amino-acids; and thus he discusses, with width, energy transformations in living matter, relating them to their underlying simpler factors.

In the same field of interest, M. Dixon & E. C. Webb deal with the phosphokinases, of which an example is hexokinase, the enzyme concerned in the first reactions of glucose. With their essential SH groups, the phosphokinases form important factors in the liberation of energy in metabolism.

An article by C. N. Hinshelwood strikes a reflective note in his application of physico-chemical theory to the problem of adaptive enzymes in bacteria; he concludes that adaptive changes in enzymic organization occur; and he discusses particularly clearly the relation of his work to that done by those who support the mutation view, showing that there is no necessary contradiction between the two theories. In considering the whole cell, several authors become concerned with the newer views (emanating from Belgium and the USA) of the organization of many important metabolic enzymes in the mitochondria. For instance G. A. Levvy shows how this must be weighed in considering β-glucuronidase, in connexion with the idea that the liver content of this enzyme may be raised in cancer. The same point arises in the account by J. N. Davidson of enzymes acting on nucleic acids; in this the reader's eyes are opened to the great advances taking place in the field of ribonucleic acids, nucleotides and nucleosides, in which A. R. Todd and colleagues on the synthetic side and his own school are playing a prominent part. It is interesting and surprising that some clinical applications are already in sight. The plasma alkaline phosphatase, to the study of which E. J. King has contributed much, has achieved a recognized clinical status as a test in jaundice and bone disease; he shows how the more academic studies of Robison and Kay led to present developments.

In conclusion mention may be made of three articles dealing with phases of the metabolism of nitrogen compounds: C. S. Hanes on transpeptidation reactions,

E. F. Gale on amino-acid decarboxylases and H. Blaschko on the enzymic oxidation of amines. The relation of these subjects to the vitamin pyridoxine, to the control of nervous tissue activity and finally to the mode by which proteins are built, provides the reader with the glimpse of fascinating future studies for biochemists.

Thirty years ago Hopkins told the writer that we had an interesting ten years ahead in biochemistry ! To one whose privilege it has been to study with advanced students the progressive knowledge of enzymes over 33 years, it has indeed been an exciting quest; but the best still lies ahead with its bright promise that these global efforts to apply chemistry and biochemistry to enzymes will prove powerful enough to establish a therapy based upon logical knowledge of the biochemistry of the cell and so to transform much that constitutes pathology today.

Rudolph Peters

We are indebted to Sir Rudolph Peters for having acted as Chairman of the committee which arranged this symposium.

Notes on Contributors

DR H. BLASCHKO is Senior Research Officer in the Department of Pharmacology at the University of Oxford. He began research in Germany in Professor O. Meyerhof's laboratory, and from there he first visited London in 1929–30 to study under Professor A. V. Hill at University College. He returned to London in 1933, and from 1934 to 1944 worked in the Physiological Laboratory at Cambridge, where he became interested in problems of amine metabolism. Dr Blaschko has worked in the Department of Pharmacology at Oxford since 1944.

PROFESSOR J. N. DAVIDSON was appointed to the Gardiner Chair of Physiological Chemistry in the University of Glasgow in 1948. He was educated at the University of Edinburgh and at the Kaiser Wilhelm Institut für Zeliphysiologie, Berlin. From 1938 to 1945 he held the post of Lecturer in Biochemistry in the University of St Andrews at University College, Dundee, and later in the University of Aberdeen. In 1945 he joined the Scientific Staff of the Medical Research Council at the National Institute for Medical Research, Hampstead, and later became Professor of Biochemistry in the University of London at St Thomas's Hospital Medical School. His early work was on fructose metabolism and enzyme biochemistry, but during the last ten years his main researches have been on the nucleic acids and on cytochemistry. From 1946 to 1952 he was Secretary of the Biochemical Society. He is a Fellow of the Royal Institute of Chemistry and of the Royal Society of Medicine and is a member of the Advisory Committee on Medical Research for Scotland appointed by the Secretary of State. In 1952 he delivered one of the four congress lectures at the Second International Congress of Biochemistry in Paris. He has published numerous papers on biochemical subjects and two books: *The biochemistry of the nucleic acids* (1950) and (in collaboration with G. H. Bell & H. Scarborough) *Textbook of physiology and biochemistry* (1950).

PROFESSOR FRANK DICKENS occupies the Philip Hill Chair of Experimental Biochemistry at the Courtauld Institute of Biochemistry, Middlesex Hospital Medical School, University of London. An article by Professor Dickens entitled " The influence of the solvent on the carcinogenic response " appeared in *British Medical Bulletin*, 1947, **4**, 348, with which a biographical note was published.

DR MALCOLM DIXON holds a Professorial Fellowship at King's College, Cambridge, and has been a member of the academic staff of the Department of Biochemistry at the University of Cambridge for 30 years. His first appointment there was in 1923 as Senior University Demonstrator, and in 1928 he became University Lecturer. Since 1945 he has been Reader in Enzyme Biochemistry and Director of the University Sub-Department of Enzyme Biochemistry. He became a Fellow of the Royal Society in 1942. He was a pupil and later a collaborator of the late Sir Frederick Gowland Hopkins. During the Second World War he directed an extra-mural research team for the Ministry of Supply, working on the enzyme biochemistry of poison gases. Dr Dixon has been interested particularly in respiratory enzymes and in manometric methods for the measurement of tissue respiration. He is the author of *Manometric methods* (3rd ed., 1952), of *Multi-enzyme systems* (1949) and of numerous papers on enzymes.

DR E. F. GALE is Director of the Medical Research Council's Unit for Chemical Microbiology at Cambridge, and is also Reader in Chemical Microbiology in the University of Cambridge. He worked out the conditions necessary for the production of amines by bacteria and, in collaboration with H. M. R. Epps, first obtained the amino-acid decarboxylases in a cell-free state; these investigations eventually led to the identification of pyridoxal phosphate as the prosthetic group of these enzymes. He is now studying the synthesis of proteins by bacteria and the effect of antibiotics on the processes involved. Dr Gale has recently been elected into the Fellowship of the Royal Society. His publications include *The chemical activities of bacteria* (1947), " Bacterial amino-acid decarboxylases " (*Advanc. Enzymol.* 1946, **6**, 1) and papers in scientific journals.

PROFESSOR C. S. HANES, after graduation from the University of Toronto in 1925, came to the University of Cambridge as an Overseas Scholar of the Exhibition of 1851, and remained there as Senior Scholar of the same Exhibition. From 1935 to 1943 he was on the staff of the Low Temperature Research Station, Cambridge, and spent the following year as scientific member of the British Food Mission in North America. Up to the war he spent most of his time on research, and his work on the mechanisms of interconversion of carbohydrates culminated in elucidation of the amylose phosphorylase of the higher plant, through which starch is synthesized. From 1944 to 1947 he was Director of Food Investigation in the Department of Scientific and Industrial Research. From then until 1951 Professor Hanes was Director of the Agricultural Research Council's Unit of Plant Biochemistry, and Reader in Plant Biochemistry at the University of Cambridge. Since 1951 he has been Professor of Biochemistry at the University of Toronto, where he is engaged on research into the mechanisms of biological synthesis, especially of peptides. In 1942 he was elected into the Fellowship of the Royal Society and in 1951 into an Honorary Fellowship of Downing College, Cambridge. From 1930 to 1938 he published a series of papers dealing with various aspects of the action of amylases, and in 1940 two papers demonstrating the reversible formation of starch from glucose-1-phosphate, catalysed by phosphorylase. Recent papers on transpeptidation and transamidation reactions are referred to in his contribution to this number of *British Medical Bulletin*.

PROFESSOR SIR CYRIL HINSHELWOOD was, from 1920 to 1921, a Fellow of Balliol College, Oxford, and for the next 16 years was a Fellow and Tutor of Trinity College, Oxford. Since 1937 he has been Dr Lee's Professor of Chemistry at the University of Oxford, and a Fellow of Exeter College. From 1946 to 1948 he was President of the Chemical Society. He is Chairman of the Fuel Research Board, and since 1950 has been Foreign Secretary of the Royal Society. Sir Cyril Hinshelwood is an honorary or foreign member of various scientific societies and academies in Europe and America.

NOTES ON CONTRIBUTORS

The Royal Society has awarded him the Davy Medal and the Royal Medal. He was awarded the Longstaff Medal and the Faraday Medal by the Chemical Society, the Lavoisier Medal by the Société Chimique de France, and the Guldberg Medal by the University of Oslo. He is the author of *Kinetics of chemical change* (1940), *Chemical kinetics of the bacterial cell* (1946), *The structure of physical chemistry* (1951), and of papers in the *Proceedings of the Royal Society* and in other scientific journals.

PROFESSOR D. KEILIN is a Fellow of Magdalene College, Cambridge. From 1925 to 1931 he was Lecturer in Parasitology at the University of Cambridge, and from 1931 to 1952 was the Quick Professor of Biology and Director of the Molteno Institute of Biology and Parasitology at the University of Cambridge. In 1928 he was made a Fellow of the Royal Society, by whom he was awarded the Royal Medal in 1939 and the Copley Medal in 1951. Professor Keilin has published papers on protozoa, insects and intracellular respiratory catalysts.

PROFESSOR E. J. KING was joint author with Dr I. D. P. Wootton and Dr J. Maclean Smith of an article on the quantitative approach to hospital biochemistry, which appeared in *British Medical Bulletin*, 1951, **7**, 307, and with which a biographical note was published.

PROFESSOR H. A. KREBS, in collaboration with Mr H. L. Kornberg, has previously contributed to the Bulletin (*Brit. med. Bull.* 1952, **8**, 206) an article on carbon dioxide exchanges between tissues and body fluids, with which a biographical note was published.

DR G. A. LEVVY is a graduate in science of the University of Edinburgh. In 1938 he joined the staff of the Pharmacology Department in the same University, and was at first an assistant and later a lecturer; in 1944 he transferred to the Biochemistry Department, where he held, from 1944 to 1945, a Beit Memorial Fellowship, and, from 1945 to 1949, an Imperial Chemical Industries Research Fellowship. In 1949 he went to the Rowett Research Institute, Aberdeen, where he is head of the Enzymology Department. Throughout most of the Second World War he was engaged in studying the metabolism of arsenic-containing compounds in the animal body, on behalf of the Ministry of Supply; this work was done in the University of Edinburgh. In 1944 he commenced a study of the significance of the enzyme β-glucuronidase in the body. Since going to the Rowett Research Institute, he has continued the study of β-glucuronidase and other enzymes concerned in uronic acid metabolism. His department has also taken up the study of cellulolytic enzymes. He has published several papers, mostly in the *Biochemical Journal*, and mainly dealing with arsenic metabolism or with β-glucuronidase and related enzymes.

PROFESSOR SIR RUDOLPH PETERS has contributed two previous articles to the Bulletin: the first, on the biochemical lesion in thermal burns, appeared in *British Medical Bulletin*, 1945, **3**, 81, in which a biographical note was published; the second paper was entitled "Development and theoretical significance of British Anti-Lewisite (BAL)" and was published in *British Medical Bulletin*, 1948, **5**, 313.

PROFESSOR J. H. QUASTEL is Professor of Biochemistry at McGill University, Montreal, and Director of the Research Institute at the Montreal General Hospital. His work concerns research in the field of biochemistry in relation to medical problems, especially cancer, antibiosis and the mode of action of drugs. In 1924 he became a Fellow of Trinity College, Cambridge, and five years later became Director of Research at Cardiff City Mental Hospital, Wales. He was Director of the Agricultural Research Council's Unit of Soil Metabolism from 1941 to 1947, after which he took up his present post in Canada. Professor Quastel was awarded the Meldola Medal by the Royal Institute of Chemistry in 1927, and became a Fellow of the Royal Society in 1940. His work since 1921 has covered aspects of the biochemistry of micro-organisms, animal tissues, plants and soils; he has published over 150 articles. He initiated biochemical work on suspensions of "resting" or "non-proliferating" micro-organisms, and on the parts played by dehydrogenase systems in the phenomenon of bacterial growth; he has also investigated the biochemistry of the brain, especially the respiratory systems, in relation to narcosis, and to aspects of mental disorder, and the mode of action of drugs on enzyme systems.

DR E. C. SLATER is working on respiratory enzyme systems at the Molteno Institute, University of Cambridge, on behalf of the Agricultural Research Council. After graduating M.Sc. from the University of Melbourne in 1939, he was appointed biochemist at the Australian Institute of Anatomy, Canberra, where he worked on nutritional problems, mainly concerned with vitamin B_1 and riboflavin. In 1946, he joined the Molteno Institute, first as a British Council Scholar (1946–48), then as an Australian National University Research Fellow (1948–51). In 1949 he visited the United States of America as a Rockefeller Fellow, working in the Department of Pharmacology, New York University. Dr Slater has published a number of papers on vitamin B_{12}, riboflavin, the succinic oxidase system, the dihydrocozymase oxidase system and oxidative phosphorylation.

PROFESSOR R. H. S. THOMPSON has previously contributed to the Bulletin (*Brit. med. Bull.* 1948, **5**, 319) a paper on therapeutic applications of British Anti-Lewisite (BAL), with which a biographical note was published. From 1938 to 1947 he was the Medical Fellow of University College, Oxford, and Demonstrator in Biochemistry, Oxford. His present work is on the cholinesterases and other enzymes of the nervous system, in relation to nerve function and nervous diseases. He has published papers in scientific journals and is joint author with C. W. Carter of *Biochemistry in relation to medicine* (1949), which was reviewed in *British Medical Bulletin*, 1950, **7**, 93.

DR E. C. WEBB of Clare College, Cambridge, has been engaged for more than ten years on teaching and research in biochemistry at the University of Cambridge. During the Second World War he worked on the biochemistry of chemical warfare agents for the Ministry of Supply and carried out early studies on the action of di-*iso*propyl fluorophosphonate (DFP) on cholinesterase. He held a Beit Memorial Research Fellowship from 1944 to 1946, when he became a University Demonstrator in Biochemistry at the University of Cambridge; in 1950 he was appointed University Lecturer. Dr Webb is a member of the Board of Governors of the Cambridge United Hospitals and from 1946 to 1952 was a member of the Cambridgeshire County Council. He is interested particularly in the purification and kinetics of enzymes, and, together with Dr K. Bailey, F.R.S., isolated yeast pyrophosphatase and crystallized yeast hexokinase.

DR D. D. WOODS has devoted his recent work to the role of folic acid and related growth factors in synthesis of amino-acids. He has previously contributed to the Bulletin (*Brit. med. Bull.* 1948, **5**, 308) a paper entitled "Wider aspects of studies of the nutrition of micro-organisms", with which a biographical note was published. Dr Woods was elected into the Fellowship of the Royal Society in 1952.

* * *

CYTOCHROME

D. KEILIN F.R.S.
E. C. SLATER Ph.D.

Molteno Institute of Biology and Parasitology
University of Cambridge

1 Myohaematin and histohaematin
2 Study of cellular respiration between 1884 and 1925
3 Cytochrome
  a Distribution of cytochrome
  b The absorption spectrum and the complex nature of cytochrome
  c Oxidation and reduction of cytochrome
4 Mechanism of indophenol reaction and cytochrome oxidase
  a Inhibition of cellular respiration by carbon monoxide and the photochemical absorption spectrum of "CO-respiratory ferment" complex
  b Cytochrome oxidase and cytochrome a_3
5 Properties of the four components of cytochrome
6 Respiratory chain
7 Oxidative phosphorylation
8 Plant cytochrome and photosynthesis
References

It is now well established that the respiration of an organism is an intracellular process. The cells have a remarkable property of oxidizing with great velocity a large variety of substances which are very stable and completely inert towards molecular oxygen. They owe this property to the presence of intracellular catalysts organized into more or less complicated systems.

The chief reaction of aerobic intracellular respiration consists in the oxidation of hydrogen atoms of metabolites by oxygen to water, the reaction which provides the energy necessary for the normal activity of the cell. The oxidation of a substrate molecule such as glucose to carbon dioxide and water consists in step-by-step dehydrogenations which are catalysed by highly specific enzymes known as dehydrogenases. This process is accompanied by a series of other important reactions such as hydrations and dehydrations, phosphorylations and dephosphorylations, and decarboxylations.

A dehydrogenase is a conjugated protein compound usually provided with a characteristic prosthetic group and a highly specific centre for attachment of the substrate molecule. Dehydrogenases combine with the substrate molecule to form a complex in which the substrate molecule undergoes activation, with the result that two of its hydrogen atoms become labile and are transferred to the prosthetic group. The substrate, or hydrogen donor, becomes oxidized while the prosthetic group, or hydrogen acceptor, becomes reduced. The reduced prosthetic group does not react with oxygen directly but transfers its hydrogen atoms (either directly or through intermediate carriers such as flavoproteins) to the cytochrome system, which reacts with oxgyen. Thus hydrogen atoms from substrate molecules all converge towards the cytochrome system, which acts as a general hydrogen acceptor reacting with molecular oxygen and reducing it not to peroxide but to water.

The cytochrome system thus forms the main if not the only path for hydrogen transfer from the activated molecules of substrates to oxygen. In other words it enables an aerobic organism to use oxgyen for its respiratory processes.

The study of the cytochrome system will be greatly helped by a brief account of the main stages in the development of our knowledge of this system in relation to intracellular respiratory activity.

1. Myohaematin and Histohaematin

Between 1884 and 1886 Mac Munn described under the names of myohaematin and histohaematin a pigment which he found in muscles and other tissues of representatives of many orders of the animal kingdom. He found that in the reduced state the pigment shows four absorption bands which are absent in its oxidized state. However, he had great difficulty in observing the actual oxidation and reduction without the help of strong oxidizing and reducing reagents such as hydrogen peroxide and ammonium sulphide. He believed that in certain tissues this pigment can co-exist with haemoglobin whereas in other tissues it can either be completely replaced by haemoglobin or it can replace haemoglobin; unfortunately he wrongly assumed that pigeon-breast muscle is completely free of haemoglobin, myohaematin being the only colouring substance of this muscle. He was able to extract myohaematin in a somewhat modified form from some tissues, including pigeon-breast muscle. Mac Munn concluded that myohaematin and histohaematin are primitive respiratory pigments on the way to develop into haemoglobin.

Mac Munn's work was very severely criticized by Hoppe-Seyler and Levy (a pupil of Hoppe-Seyler) who rightly showed that pigeon-breast muscle does contain haemoglobin (myoglobin) and ascribed the four-banded spectrum observed by Mac Munn to a mixture of oxyhaemoglobin, haemoglobin and a little haemochromogen. After a brief polemical discussion between Mac Munn and Hoppe-Seyler the authority of the latter prevailed and myohaematin or histohaematin disappeared from scientific literature for more than 40 years. The pigment never reached textbooks and when occasionally it did emerge from oblivion it was only to illustrate the correctness of Hoppe-Seyler's criticism and later it was completely confused by Günther with myoglobin or muscle haemoglobin.

2. Study of Cellular Respiration between 1884 and 1925

While the important work of Mac Munn was rejected and completely forgotten, the interest in cell respiration was steadily growing. We shall briefly summarize the developments of only certain aspects of this work, which have a direct bearing on our problem and which have taken place between 1884, the year of the discovery of myohaematin or histohaematin, and 1925, the year of its rediscovery and the reintroduction of haematin compounds as intracellular respiratory catalysts.

The concept originating from the work by Spallanzani (1807) that the respiration of an organism is localized not in its blood but within its tissues had a long and complicated course and was completed only by Pflüger (1875). However, another 20 years passed before this concept was consolidated and generally accepted.

The study of haemoglobin and its derivatives, already well advanced in Mac Munn's time, through the work of Hoppe-Seyler, Stokes and others, rapidly developed. The reactions of haemoglobin with O_2 and with CO were carefully investigated. The photodissociation of CO-haemoglobin was demonstrated by Haldane & Smith (1896). The true nature of methaemoglobin was clarified by Conant (1923) and Conant & Fieser (1925) and its reactions with cyanide and fluoride investigated. Haemochromogen was finally demonstrated as a compound of a reduced haematin or haem with nitrogenous substances. The property of combining reversibly with CO was found to be shared by all ferrous derivatives of haematin, such as haemoglobin, haemochromogen and haem. The chemical structure of porphyrin was finally elucidated.

The study of biological oxidation reactions opened by Planche, Schönbein, Hoppe-Seyler and Traube was actively pursued by other workers. Bertrand (1897), in the course of his study of the mechanism of darkening of the latex of lacquer trees, recognized that this phenomenon is caused by the oxidation of phenolic substances like urishiol or laccol catalysed by an enzyme described by him under the name of laccase. For this and another enzyme named by him tyrosinase he introduced the general term of oxidase. Having found in his enzyme preparation an appreciable concentration of manganese, he put forward the view that manganese formed part of the enzyme and was essential for its catalytic activity.

The study of oxidizing mechanisms in animal tissues originated from the work of Ehrlich (1885) on the oxygen requirements of tissues. He found that on injecting rabbits with a mixture of dimethyl-*p*-phenylenediamine and α-naphthol an oxidative synthesis of indophenol takes place in tissues of different organs. However, he never ascribed this reaction to the catalytic activity of an oxidizing enzyme. This concept is due to Röhman & Spitzer (1895), who studied the reaction in different tissues and extracts. Further work by Spitzer (1897) led him to the conclusion that the enzyme catalysing the synthesis of indophenol is an iron-nucleoprotein compound or more simply an " organically bound iron which mediates the oxygen transport." The work of Spitzer was followed by a general survey by many workers of the distribution of the indophenol reaction in all organisms from bacteria and fungi to higher plants and animals. Of special importance was an extensive investigation by Battelli & Stern (1912) of the oxidation of *p*-phenylenediamine and of succinic acid by different animal tissues and their extracts. They followed the oxidation of both substances by the oxygen uptake and found that the activity of the oxidase catalysing these reactions is proportional to the respiratory activity of the corresponding tissue. Working with disintegrated cells and tissues they found that the enzyme, unlike other known oxidases, could not be obtained in solution. They introduced the term " oxydones " for insoluble enzymes which they found were linked with the main respiratory activity of the cell and were reversibly inhibited by cyanide and narcotics.

Parallel with the work on oxidizing enzymes two rival theories of the mechanism of biological oxidation reactions were developed, one by Wieland and Thunberg, the other by Warburg.

1. According to Wieland (1922), the biological oxidation reactions are catalysed by enzymes, called by him dehydrases but generally known as dehydrogenases, which activate certain hydrogen atoms of substrate molecules with the result that they become labile and can be transferred to a suitable hydrogen acceptor such as quinone, methylene blue or oxygen. Thus he found it unnecessary to postulate the activation of oxygen, which acts merely as an acceptor of activated hydrogens. According to this theory, which was greatly extended by Thunberg (see review by Thunberg, 1937), dehydrogenases are highly specific, a different enzyme being required for each substrate.

2. Warburg's view on the mechanism of biological oxidation reactions was well summarized in the paper read by him at the meeting of the German Chemical Society held on 29 April 1925. According to Warburg (1925), all phenomena connected with cellular respiration can be brought down to the action of two sorts of forces: the unspecific surface forces of the solid components of the cell and the specific chemical forces—the respiratory ferment and its oxygen-transporting component, the iron. The iron nature of the respiratory ferment was believed by him to be supported by the following evidence :

i. all organisms or cells contain iron, the presence of which is essential for their normal life;
ii. the total iron content of cells and tissues is sufficient to cope with their normal respiratory activity, i.e., the rate of their oxygen uptake per milligram of total iron is comparable with the catalytic activity of iron in the oxidation of cysteine;
iii. the addition of a small amount of iron to the granular extract of unfertilized sea-urchin eggs increases its oxygen uptake;
iv. a low concentration of cyanide inhibits both respiration and oxidation reactions catalysed by iron, in vitro;
v. respiration can to some extent be imitated by oxidation of certain amino-acids catalysed by iron-containing charcoals prepared from cane sugar, blood, haemin or different commercial dyes such as safranin, neutral red or Bismarck brown, the activity of which is also inhibited by cyanide and narcotics;
vi. the oxidation of unsaturated fatty acids and of cysteine is strongly catalysed by iron.

Warburg concluded that the " respiratory ferment " is the sum of all the catalytically active iron compounds present in the cell. Although the results of his work between 1908 and 1925 strongly support the view that metal catalysis is involved in cellular respiration, the nature of the metal could not be ascertained. The suggestion put forward by Spitzer (1897) and by Warburg during the period 1908 to 1925 (see Warburg, 1925) that the metal in question is iron could not be definitely demonstrated.

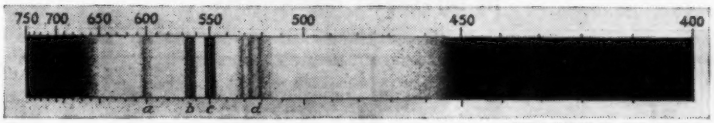
In spite of an overwhelming weight of evidence accumulated at that time pointing to the important part played in respiration by the highly specific enzymes—the dehydrogenases—Warburg strongly defended his theory that cellular respiration consists in the oxidation of metabolites condensed upon the unspecific intracellular surfaces by an iron-containing " respiratory " or " oxygen-transporting ferment ".

In 1924, Fleisch and Szent-Györgyi, in the course of their study of the oxidation of *p*-phenylenediamine and succinic acid by different tissues, made independent attempts to reconcile the two rival theories of respiration (Fleisch, 1924; Szent-Györgyi, 1924) but their work failed to convince the supporters of these two theories.

It may be mentioned here that during this period the literature on haemoglobin and its derivatives on the one hand, and on biological oxidation reactions on the other, grew at a very rapid pace and prior to 1925 several extensive

reviews of many hundreds of papers on different aspects of this subject were published by different workers. However, none contained any reference to Mac Munn's work. It is therefore not surprising that neither Warburg, who between 1908 and 1925 showed great interest in metal catalysis, nor the other workers interested in biological oxidation reactions, came across Mac Munn's work. In fact nothing short of the rediscovery of this pigment could bring his work to light.

FIG. 1. ABSORPTION SPECTRUM OF REDUCED CYTOCHROME



Keilin (1925), by permission of the Royal Society

Scale: wavelength (mμ.)

a: α-band of cytochrome *a*
b: α-band of cytochrome *b*
c: α-band of cytochrome *c*
d: fused β-bands of at least cytochromes *b* and *c*

3. Cytochrome

While the respiration of *Gastrophilus* larva and especially the fate of its haemoglobin in the pupal and adult stages of this insect were being studied, a pigment was discovered which was identified with Mac Munn's myohaematin or histohaematin. In the course of this work (Keilin, 1925), not only were Mac Munn's observations corroborated but the pigment was found to be much more widely distributed in nature and of a much greater physiological importance than was anticipated by Mac Munn. The pigment was described under the name of cytochrome, which means cellular pigment, for the reason which will be more apparent after enumeration of the main points in which the new findings concerning it differed from those of Mac Munn:

1. The pigment is present not only in tissues of animals but also in those of plants, as well as in unicellular organisms such as yeasts and bacteria.
2. The multibanded absorption spectrum of this pigment, as seen by direct spectroscopic examination of cells, revealed that it belongs not to one substance but to three haematin compounds showing marked differences in structure and properties.
3. Each of the components is not a free haematin but a haematin bound to a nitrogenous substance, the protein; hence the term haematin applied to them is unsuitable.
4. The pigment cannot replace or be replaced by haemoglobin, since the two have entirely different functions, and it cannot therefore be considered as an immature pigment on the way to develop into haemoglobin.
5. The fact that the pigment can be seen in the uniform suspension of living intact yeast cells and bacteria made it possible:
   i. to establish beyond doubt its intracellular nature;
   ii. to study directly by means of a low-dispersion spectroscope its oxidation and reduction in living intact cells and to correlate its activity with the respiratory activity of cells;
   iii. to study the effects and to localize the action of respiratory inhibitors such as narcotics and cyanide.
6. The study of cytochrome made it possible to correlate its activity with that of other intracellular respiratory catalysts and to reconcile the two rival theories of respiration, that of Wieland and Thunberg and that of Warburg, in showing that they were merely dealing with two portions of the same respiratory system of the cell.

a. Distribution of Cytochrome

Cytochrome is very widely if not universally distributed in cells of aerobic organisms—from bacteria and yeast to higher plants and animals. It is absent from anaerobic organisms such as strictly anaerobic bacteria and bottom yeast. Moreover there is in nature a certain parallelism between the concentration of this pigment and the respiratory activity or oxygen requirements of cells. Thus the highest concentration of cytochrome is found in muscles capable of rhythmic or rapid contractions, such as heart muscle of higher animals, wing muscles of flying insects and pectoral muscles of flying birds, whereas the pectoral muscles of the domestic fowl contain very little cytochrome. A high concentration of cytochrome is also found in the grey matter of brain, which is very sensitive to anoxia. Vigorously respiring cells of *Torula utilis*, of baker's yeast or of strictly aerobic bacteria like *Bacillus subtilis* are also very rich in cytochrome.

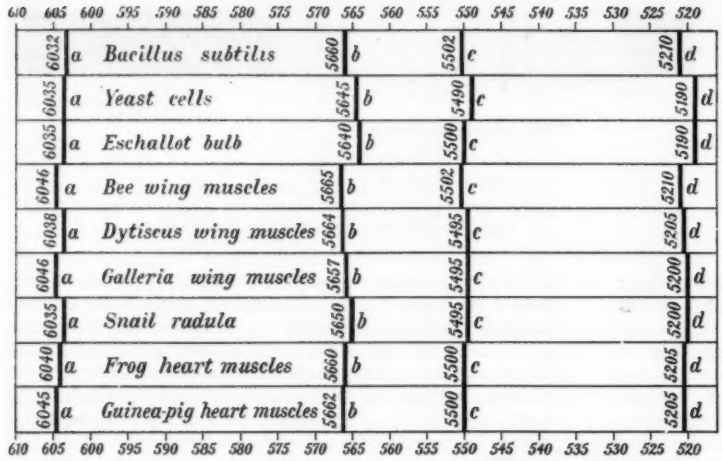
b. The Absorption Spectrum and the Complex Nature of Cytochrome

In the reduced state, cytochrome shows four distinct absorption bands (fig. 1) occupying approximately the following positions: a: 605 mμ.; b: 562 mμ.; c: 550 mμ. and d: 520 mμ. The positions of these bands in a typical cytochrome show very little variation from one organism to another (fig. 2). This complex absorption spectrum belongs not to one compound but to three haemochromogen-like components denoted as cytochrome *a*, cytochrome *b* and cytochrome *c*. The bands a, b and c represent the α-bands of these three haemochromogens whereas band d corresponds to the fused β-bands of at least two components: cytochrome *b* and *c*. To this point as well as to the structure of these components we shall return later.

c. Oxidation and Reduction of Cytochrome

As cytochrome in the reduced state shows intense absorption bands of a haemochromogen type which on oxidation are replaced by diffuse hardly perceptible bands of a parahaematin type, the oxidation and reduction of cytochrome can be easily observed in intact living cells such as a suspension of yeast or bacteria, or in the wing muscles of a living moth (*Galleria*). When a suspension of yeast is examined with a low-dispersion hand spectroscope or microspectroscope, the absorption bands described above are clearly visible. The absorption bands disappear when air is passed through the suspension, and rapidly reappear when this stands. The rate of oxidation and reduction of cytochrome reflects the degree of the respiratory activity of cells. Saturation of cells with substrate or metabolites accelerates the reduction of cytochrome whereas removal of substrate or starvation slows it down. Narcotics such as urethane or alcohols slow down the reduction of components *a* and *c*, but prevent the oxidation

FIG. 2. ABSORPTION SPECTRUM OF REDUCED CYTOCHROME IN DIFFERENT ORGANISMS



Based on figures of Keilin (1925), by permission of the Royal Society

Scale: wavelength (mμ.)

a: α-band of cytochrome *a*
b: α-band of cytochrome *b*
c: α-band of cytochrome *c*
d: fused β-bands of cytochromes *b* and *c*

Figures on the bands are positions of maximum absorption in A. (1 A. = 0.1 mμ.)

of component *b*. Addition of cyanide even in a very small concentration reversibly inhibits the reoxidation of reduced cytochrome.

The state of cytochrome as seen spectroscopically reflects only the difference between the rates of its oxidation and reduction. Thus, in the thoracic muscle of a living moth (*Galleria*), the completely oxidized cytochrome becomes at once reduced when a current of air containing a little hydrogen cyanide is passed over the moth; the cyanide does not reduce cytochrome but only prevents its reoxidation.

The study of the mechanism of the oxidation and reduction of cytochrome was carried out on yeast cells treated in different ways and on a cell-free heart-muscle preparation which was free of haemoglobin and myoglobin but very rich in cytochrome, succinic dehydrogenase and certain other enzymes (Keilin, 1927a, 1927b, 1929).

The fact that the rate of reduction of cytochrome was found to depend upon the presence of a suitable substrate, and to be affected in the same way and to the same degree as the rate of reduction of methylene blue by the same preparation, clearly indicated that cytochrome is reduced by the dehydrogenase-substrate systems.

The fact that the oxidation of cytochrome was found to be strongly and reversibly inhibited by cyanide without affecting the absorption spectra of its three components (*a*, *b* and *c*) clearly indicated that these components are not autoxidizable and that their oxidation is brought about by an oxidizing enzyme or oxidase. Among the oxidases already well known, the indophenol oxidase fulfilled the necessary requirements for such a catalyst since its distribution is almost as wide as that of cytochrome, its activity is proportional to the respiratory activity of cells, and it is reversibly inhibited by a low concentration of cyanide. Moreover indophenol oxidase was found to show a light-sensitive inhibition by CO in every respect similar to that of the respiration of intact cells which was demonstrated by Warburg and his co-workers from 1926 to 1929, and which will be examined later on. Thus, indophenol oxidase was identified with Warburg's respiratory ferment (Keilin, 1927a, 1927b, 1929).

Cytochrome was therefore found to occupy a central position linking the dehydrogenase systems with indophenol oxidase. The hydrogen atoms of the substrate molecules activated by specific dehydrogenases are transferred through cytochrome and the indophenol oxidase to molecular oxygen, the activity of cytochrome in this respiratory system being due to valency changes of its iron atoms.

4. Mechanism of Indophenol Reaction and Cytochrome Oxidase

Indophenol oxidase shares with phenol oxidases the property of being reversibly inhibited by cyanide and by carbon monoxide. However, whereas indophenol oxidase is insoluble or intimately linked with the structural elements of the cell and its inhibition by CO is light-sensitive, phenol oxidase is perfectly soluble and its inhibition by CO is light-insensitive (Keilin, 1929). The insolubility of the indophenol oxidase, which precluded its isolation and purification, made impossible the study of its activity independently of cytochrome components. In fact, the most active preparation of indophenol oxidase was represented by a colloidal heart-muscle preparation containing the cytochrome system. The isolation and purification of cytochrome *c* from baker's yeast (Keilin, 1930) and later by a simpler method from heart-muscle (Keilin & Hartree, 1937) greatly facilitated the study of indophenol oxidase. In the first instance it was found that neither the heart-muscle preparation itself nor cytochrome *c* catalyses the oxidation of cysteine to cystine, but when the two are brought together they form a strong catalytic system for this oxidation. The reaction was completely inhibited by cyanide and showed a marked light-sensitive inhibition by CO (Keilin, 1930). Although cytochrome *c* is present in the heart-muscle preparation, it appears to be inaccessible to cysteine. Similar results were later obtained with hydroquinone, ascorbic acid, adrenaline and other reducing agents in place of cysteine (Keilin & Hartree, 1938; see also Slater, 1949a). *p*-Phenylenediamine, on the other hand, readily reduces the oxidized cytochrome *c* in the oxidase preparation, so that it is oxidized in the absence of added cytochrome *c*. Even with this substance, however, the oxidation was augmented by the addition of cytochrome *c*. All these reducing agents are therefore not oxidized directly by the oxidase but through co-operation of cytochrome. The only catalytic property which can be definitely ascribed to the oxidase is the oxidation of reduced cytochrome. The correct name for this enzyme should therefore be cytochrome oxidase. It can be said now that the indophenol reaction is catalysed by the

" indophenol oxidase system " composed of cytochrome oxidase and cytochrome *c*.

Before we deal with the nature of cytochrome oxidase and the structure of the cytochrome components we shall briefly examine some of the results of the work carried out by Warburg and his co-workers on the inhibition of respiration by carbon monoxide.

a. Inhibition of Cellular Respiration by Carbon Monoxide and the Photochemical Absorption Spectrum of " CO-Respiratory Ferment " Complex

Of the different aspects of biological oxidation reactions investigated by Warburg up to 1925, it is mainly his study of the effect of cyanide that threw some light on the mechanism of cellular respiration. Although it brought strong evidence in support of the view that metal catalysis is involved in this process, that the metal is iron could not be definitely demonstrated.

The work on cytochrome (Keilin, 1925) summarized above, which revealed that yeast cells contain three haemochromogen-like components of cytochrome, together with the fact that all the ferrous derivatives of haemoglobin have a high affinity for CO, suggested to Warburg the study of the effect of CO on the respiratory activity of yeast cells (Warburg, 1926). He found that the respiration of yeast is inhibited by CO (but only when the partial pressure of this gas is very high) and that the degree of inhibition depends upon the ratio $\frac{CO}{O_2}$. He also found that the inhibition of respiration in the dark was reversed when yeast cells were exposed to light. Warburg concluded that CO competes with O_2 for the " respiratory ferment " which appeared to be related to haemoglobin in view of the fact that CO-haemoglobin was shown by Haldane & Smith (1896) to be a light-sensitive complex. However, the affinity of this ferment for CO was found to be between 1,500 and 3,000 times lower than that of a mammalian haemoglobin. From the degree of reactivation of the CO-inhibited respiration of yeast cells (*Torula utilis*) when exposed to light of different wavelengths and constant energy, Warburg & Negelein (1929) obtained a photochemical absorption spectrum very similar to that of a haematin compound with the bands α, β and γ lying at approximately 590 mμ., 540 mμ. and 432 mμ. The photochemical spectrum of *Torula utilis* and *Acetobacter pasteurianum* was obtained more accurately by Kubowitz & Haas (1932) using for this purpose 30 different wavelengths ranging from 253 mμ. to 671 mμ. In addition to these two spectra, the photochemical absorption spectra were obtained for only three other types of cells or tissues and they all show very marked differences in the number, the relative heights and the positions of their absorption bands. The very limited use which has so far been made of the photochemical method since its introduction more than 20 years ago is due not only to its elaborate technique but also to the fact that the respiration of cells and tissues, even when it depends upon the cytochrome system, is not always inhibited appreciably by CO, and when a marked inhibition by CO is obtained it is not always reversed by light. The significance of the photochemical absorption spectra and their interpretation was recently re-examined (Keilin & Hartree, 1953) and it was suggested that they belong not only to the CO complex of cytochrome oxidase but reflect also the absorption bands of certain other components of cytochrome which by themselves do not react with CO.

b. Cytochrome Oxidase and Cytochrome a_3

In the early period of the work on cytochrome, workers engaged in the systematic study of cytochrome in micro-organisms (Yaoi & Tamiya, 1928; Tamiya & Yamagutchi, 1933) found that in certain bacteria the component *a* of cytochrome may be replaced by components a_1, lying at about 589–590 mμ., or a_2 at about 632 mμ. In some micro-organisms both of these components may be present simultaneously. These components react either with HCN or CO or with both inhibitors and therefore can be considered as corresponding to cytochrome oxidases even if their complex with CO does not dissociate in light (e.g., *Bacterium coli* or *Azotobacter chroococcum*) and therefore does not give the photochemical absorption spectrum.

We may now ask whether it is possible to see within cells or tissues containing a typical cytochrome system (comprising the components *a*, *b* and *c* of cytochrome) an additional component easily autoxidizable and reacting with the respiratory inhibitors. On careful examination of material rich in cytochrome *a*, it was demonstrated that the bands α at 605 mμ. and γ at 450 mμ. belong not, as was hitherto believed, to one component *a* but to two components, *a* and a_3, the absorption bands of which are superimposed (Keilin & Hartree, 1939).

The component a_3 differs from other cytochrome components (i) by being much more easily autoxidizable, (ii) by combining in the ferric state with cyanide, hydrosulphide and azide, and (iii) by combining in the ferrous state with carbon monoxide, forming a complex with absorption bands at 590 mμ. and 432 mμ., occupying thus the positions of the corresponding bands in the photochemical absorption spectrum. For this reason, it was suggested that the component a_3 is cytochrome oxidase or the " oxygen-transporting ferment " of Warburg.

5. Properties of the Four Components of Cytochrome

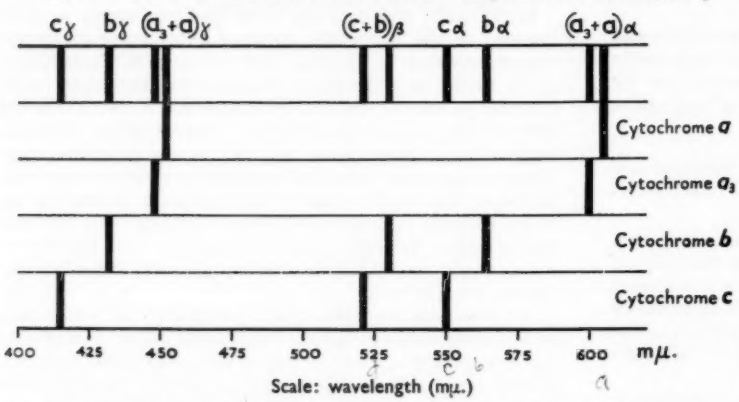
The absorption spectra of haematin compounds and of all their derivatives with closed porphyrin rings show in addition to the bands in the visible region a band in the violet region of the spectrum. This absorption band, which is by far the strongest, is known as the Soret or γ-band.

The complex multibanded absorption spectrum of all the four components of cytochrome (*a*, a_3, *b* and *c*) can be represented schematically as shown in fig. 3, the components *a* and a_3 being represented each by two bands α and γ and the components *b* and *c* by three bands α, β and γ. Each of the components is a haemoprotein compound, i.e. an iron-porphyrin protein compound.

We shall now briefly examine some of the properties of these components.

The absorption bands of components *a* and a_3 are not seen separately as shown in fig. 3 but are completely fused or superimposed (α: 603 mμ.; γ: 450 mμ.). It was only by applying to the study of these components the action of some inhibitors that the existence of the component a_3 was revealed. These two components have probably an identical protein and prosthetic group but differ only in the mode of linkage between them with the result that cytochrome a_3, unlike cytochrome *a* or other components, is autoxidizable and reacts with the respiratory inhibitors. Components *a* and a_3 are of greenish colour, their haem nucleus resembling that of chlorocruorin, which differs from protohaem of haemoglobin

FIG. 3. COMPONENTS OF CYTOCHROME (Keilin & Hartree, 1939)


cγ bγ (a3+a)γ (c+b)β cα bα (a3+a)α
Cytochrome a
Cytochrome a3
Cytochrome b
Cytochrome c
400 425 450 475 500 525 550 575 600 mμ.
Scale: wavelength (mμ.)

The individual absorption bands of cytochromes *a*, a_3, *b* and *c* are combined in the top multibanded spectrum.

in having one of the vinyl groups replaced by a formyl group and probably some additional groups the nature and localization of which are not yet determined (Rawlinson & Hale, 1949; Warburg & Gewitz, 1952). Both *a* and a_3 are thermolabile and very sensitive to acid, alkali, and organic solvents. They are intimately linked to the structural elements of cells and so far have not been completely extracted or purified to any considerable extent.

Component *b* is also a thermolabile haemoprotein compound intimately bound to the structural elements of cells, and has not therefore been isolated and purified. In the reduced state it shows three absorption bands: α: 562 mμ.; β: 530 mμ.; γ: 432 mμ. It appears to be closely linked with succinic dehydrogenase. Within the physiological range of pH it does not react with respiratory inhibitors like cyanide, azide or carbon monoxide. Outside this range it undergoes irreversible denaturation.

Cytochrome *c* on the contrary is thermostable, very resistant to alkali and acids and can easily be extracted and obtained in a pure state. It shows three absorption bands: α: 550 mμ.; β: 521 mμ.; γ: 415 mμ. It is a chromoprotein of small molecular weight of about 13,000 containing 0.43% iron. Its haem is linked with the protein both through the porphyrin and the iron. Two cysteines of the protein are supposed to form thio-ether linkages with the two vinyl groups of the porphyrin (Theorell & Åkesson, 1941). Cytochrome *c* is not autoxidizable within the physiological range of pH; only at pH above 13 and below 4 does it become autoxidizable and react with CO. In the oxidized state, cytochrome *c*, like other parahaematin compounds, is of brownish-red colour (α: 566 mμ.; β: 529 mμ.; and γ: 407 mμ.); on reduction it becomes salmon-pink. The absorption spectrum of purified cytochrome *c* is shown in fig. 4. Its physico-chemical properties were extensively investigated by Theorell & Åkesson (1941) and by Paul (1951).

6. Respiratory Chain

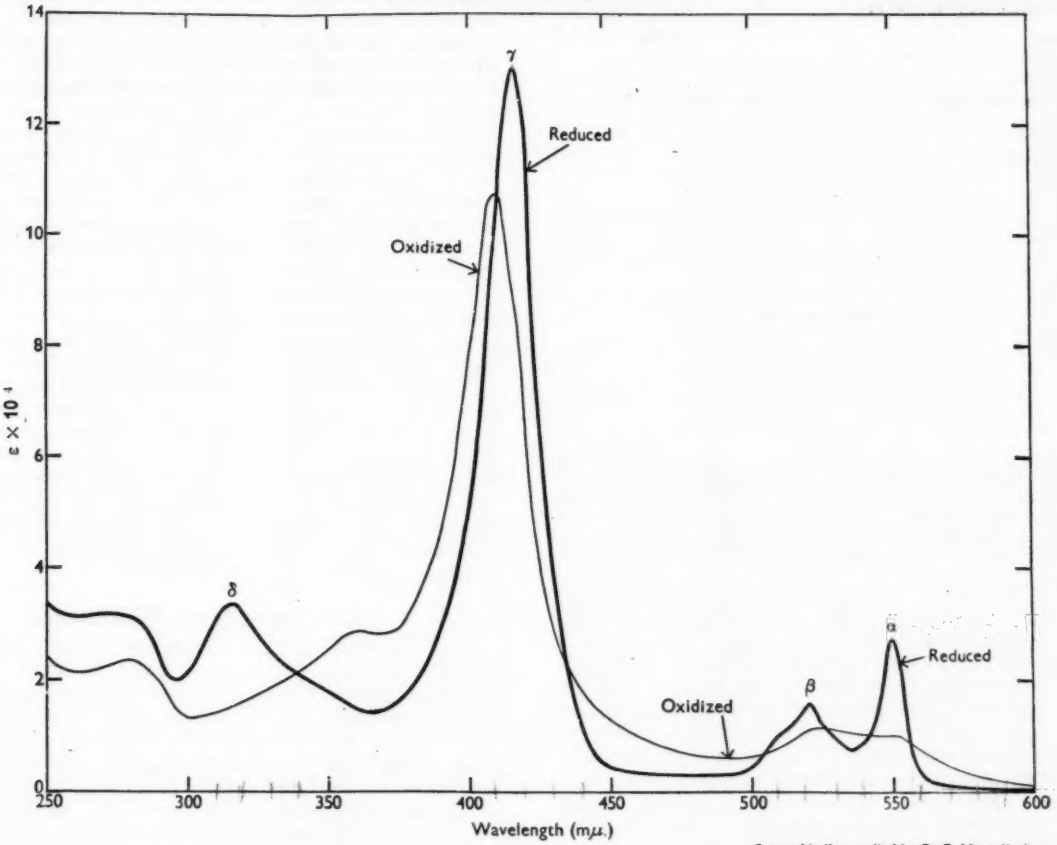
The discovery of the role of cytochrome as a link between the dehydrogenases and oxygen established the concept that a respiratory chain of intracellular catalysts is required for the aerobic oxidation of intermediary metabolites. Between 1934 and 1937, it became clear, chiefly from the work of Warburg, that many substrates require, in addition, the pyridine nucleotide coenzymes (first identified as a coenzyme of alcoholic fermentation by Harden & Young in 1906) and flavoproteins. During the last 15 years, attention has been given to the elaboration of the details of the respiratory chain. Ball (1938) measured the oxidation-reduction potentials of cytochromes *a*, *b* and *c*. A reaction scheme which describes the oxidation of succinate and of the substrates which react with diphosphopyridine nucleotide (DPN)[1] is given in fig. 5. Additional coenzymes (α-lipoic acid and coenzyme A) are involved in the oxidation of the α-keto acids, pyruvate and α-ketoglutarate.

The participation of the " factor " (shown in fig. 5) was indicated by a study of the mechanism of action of certain reducing agents (glutathione, 2 : 3-dimercaptopropanol, ascorbic acid) which, when incubated with tissue preparations in the presence of air, inactivate the respiratory chain without having any effect on the activities of dehydrogenases, coenzymes, flavoproteins, the known cytochrome components or cytochrome oxidase (Slater, 1949a, 1950a). Since haematin compounds (e.g., haemoglobin) are destroyed by coupled oxidation with reducing agents in the presence of air, it is possible that this factor is also a haematin compound. The total haematin content of a heart-muscle preparation is decreased after incubation with these reducing agents in the presence of air, although the intensities of the cytochrome bands are not affected. It has been known for some time (Keilin, 1927a) that tissues contain much more haematin than can be accounted for by the cytochromes and other identified haematin compounds. According to Potter & Reif (1952) and Chance (1952), antimycin inhibits respiration by combining stoicheiometrically with the same factor, and they have used antimycin to " titrate " the concentration of the factor. Certain other substances (narcotics (Keilin, 1925; Keilin & Hartree, 1940); certain hydroxynaphthoquinones (Ball, Anfinsen & Cooper, 1947); 4 : 4′-dihydroxystilbene (Case & Dickens, 1948); surface-active compounds (Slater, 1949c)) inhibit respiration in a manner resembling the inactivation brought about by incubation with reducing agents, but it is not certain whether these substances act specifically on the factor or inactivate the respiratory system by a non-specific physical action on the particles which contain all the components of the respiratory chain. A change in the physical properties of the particles can affect the mutual accessibility of the components of the system (Keilin & Hartree, 1949).

Battelli & Stern (1912) first showed that the main respiration of the cell was associated with insoluble fractions of the cell. Warburg (1913) also found that the respiration was carried out by the granular components of the liver cell. The heart-muscle preparation used in studies of cytochrome oxidase (Keilin, 1929; Keilin & Hartree, 1938, 1939) and of the oxidation of succinate (Keilin & Hartree, 1940, 1949) is a

[1] In view of the varying terminology for these coenzymes, it is pointed out that cozymase, coenzyme I and diphosphopyridine nucleotide are synonymous with diphospho-nicotinamide-adenine dinucleotide (DPN); while coenzyme II, triphosphopyridine nucleotide and triphospho-nicotinamide-adenine dinucleotide (TPN) also refer to one and the same coenzyme.—ED.

FIG. 4. ABSORPTION SPECTRA OF OXIDIZED AND OF REDUCED CYTOCHROME *c*, CONTAINING 0.45% IRON



Curves kindly supplied by Dr E. Margoliash (Molteno Institute, Cambridge)

Abscissae: wavelength (mμ.)

Ordinates: absorption coefficient, $\varepsilon \times 10^{-4}$

where $\varepsilon = \frac{1}{c.d} \log \frac{I_0}{I}$

I_0 and I are the intensities of the incident and transmitted light respectively

c = concentration of cytochrome (g. atoms Fe/l.)

d = optical path (cm.)

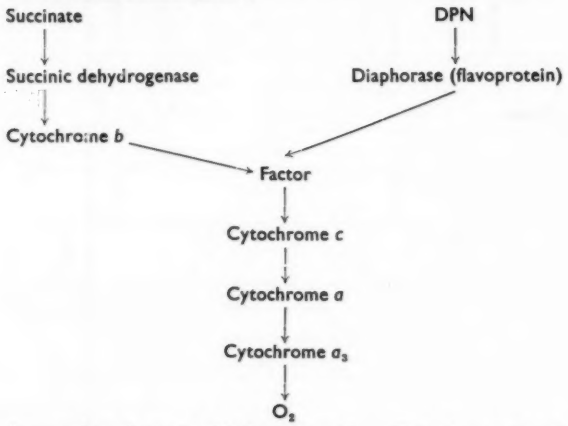
α, β, γ, δ: absorption bands

colloidal solution obtained by the disintegration of water-washed heart muscle. During the last six years, it has become recognized, chiefly owing to the work of Claude, Hogeboom, Schneider, Green and Lehninger in the USA, that most of the respiratory enzymes are situated in the mitochondria of liver and kidney (see reviews by Lehninger (1951) and Green (1951)). In heart muscle, the respiratory enzymes are localized in granules called sarcosomes, which have been long known to cytologists and are arranged in a regular manner in the sarcoplasm between myofibrils, one sarcosome being found opposite each anisotropic band of the myofibril (Cleland & Slater, unpublished). Thus, the sarcosomes correspond to mitochondria in other tissues (see also Harman & Feigelson, 1952). There is evidence that the respiratory chain is situated in a semi-permeable membrane which bounds the sarcosome and that the colloidal preparation used in studies of the oxidase systems is derived from the disintegration of this membrane; it is free from many of the more soluble and more labile components of the intact mitochondria (Cleland & Slater, 1953).

7. Oxidative Phosphorylation

The biological function of respiration is to provide the energy for those energy-consuming processes, such as synthesis of biological materials, and osmotic or mechanical work, essential for the life of the organism. In all these processes, the immediate source of the energy is probably the energy-rich terminal phosphate group of adenosinetriphos-

FIG. 5. MECHANISM OF OXIDATION OF SUCCINATE AND OF REDUCED DIPHOSPHOPYRIDINE NUCLEOTIDE (DPN)



Arrows show direction of hydrogen transfer. This scheme is based on that of Slater (1949b, 1950a), modified by taking into account the demonstration of Tsou (1951) that succinic dehydrogenase is in the reaction chain. Chance (1952) has given quantitative support to many of the reactions of this scheme, but believes that cytochrome *b* is not in the main reaction chain for the oxidation of succinate.

phate (ATP) and the chief source of the ATP is respiration.[2] That is to say, coupled with the hydrogen-transferring reactions shown in fig. 5, ATP is synthesized from adenosine-diphosphate (ADP) and inorganic phosphate. Lehninger (1949) has shown that probably three molecules of ATP are synthesized during the oxidation of one molecule of reduced DPN. Since it appears unlikely that more than one molecule of ATP can be synthesized at a time, the biological necessity of a series of hydrogen-transferring catalysts acting as a respiratory chain is clear. However, the mechanism of even one of these reactions remains completely obscure and this is one of the most important problems for the future.

[2] See Krebs, p. 97 of this Bulletin.

Little success has yet attended attempts to locate those reactions of the respiratory chain that are associated with phosphorylation. With α-ketoglutarate as substrate, nearly as high a yield of phosphorylation was obtained with cytochrome *c* as hydrogen acceptor as with oxygen (Slater, 1950b). It was suggested, therefore, that all the phosphorylation steps occurred between substrate and cytochrome *c* and none between cytochrome *c* and oxygen. This conclusion has been disputed, because so much of the energy of the respiratory chain is found between cytochrome *c* and oxygen. However, phosphorylation associated with the cytochrome oxidase reaction has not yet been directly demonstrated. It is possible that the energy liberated in the cytochrome oxidase reaction is, under certain circumstances, made available in some other way. For example, several authors have suggested that ion transport is linked with the electron transport per se, without the intervention of phosphorylation. However, this has been made less likely by the most recent work (e.g., see Davies & Krebs, 1952).

One of the most puzzling features of oxidative phosphorylation is that it is possible to carry out the reactions described in fig. 5 in the complete absence of inorganic phosphate (Keilin & Hartree, 1949; Bonner, 1951). Moreover, although heart-muscle sarcosomes are able to couple these reactions with phosphorylation, the preparation of disintegrated sarcosomes promotes the very rapid oxidation of succinate and reduced DPN without any phosphorylation even when inorganic phosphate and ADP are added.

8. Plant Cytochrome and Photosynthesis

Hill & Scarisbrick (1951) and Davenport & Hill (1952) have extensively studied the cytochrome system in plants. Plants contain, as well as a normal cytochrome system, additional cytochromes named cytochrome b_3 and *f*. Cytochrome *f* resembles cytochrome *c* but there are important differences; it has a much higher oxidation-reduction potential and is confined to the green parts of the plant. Davenport & Hill (1952) have suggested that cytochrome *f* may be reduced in the photosynthetic reaction, with the evolution of oxygen, i.e., the role of cytochrome *f* is the opposite of that of the other cytochromes, which are concerned in the uptake of oxygen.

REFERENCES

Ball, E. G. (1938) *Biochem. Z.* **295**, 262
Ball, E. G., Anfinsen, C. B. & Cooper, O. (1947) *J. biol. Chem.* **168**, 257
Battelli, F. & Stern, L. (1912) *Ergebn. Physiol.* **15**, 96
Bertrand, G. (1897) *C.R. Acad. Sci., Paris*, **124**, 1032; 1355
Bonner, W. D. (1951) *Biochem. J.* **49**, *Proc.* viii
Case, E. M. & Dickens, F. (1948) *Biochem. J.* **43**, 481
Chance, B. (1952) *Nature, Lond.* **169**, 215
Cleland, K. W. & Slater, E. C. (1953) *Biochem. J.* **53**, 547
Conant, J. B. (1923) *J. biol. Chem.* **57**, 401
Conant, J. B. & Fieser, L. F. (1925) *J. biol. Chem.* **62**, 595
Davenport, H. E. & Hill, R. (1952) *Proc. roy. Soc.* B, **139**, 327
Davies, R. E. & Krebs, H. A. (1952) *Symp. biochem. Soc.* No. 8, p. 77
Ehrlich, P. (1885) *Das Sauerstoff-Bedürfniss der Organismus. Eine farbenanalytische Studie.* Hirschwald, Berlin
Fleisch, A. (1924) *Biochem. J.* **18**, 294
Green, D. E. (1951) *Biol. Rev.* **26**, 410
Haldane, J. & Smith, J. L. (1896) *J. Physiol.* **20**, 497
Harden, A. & Young, W. J. (1906) *Proc. roy. Soc.* B, **77**, 405
Harman, J. W. & Feigelson, M. (1952) *Exp. Cell Res.* **3**, 47; 508
Hill, R. & Scarisbrick, R. (1951) *New Phytologist*, **50**, 98
Keilin, D. (1925) *Proc. roy. Soc.* B, **98**, 312
Keilin, D. (1927a) *C.R. Soc. Biol., Paris*, **97**, Suppl., p. 39
Keilin, D. (1927b) *Nature, Lond.* **119**, 670
Keilin, D. (1929) *Proc. roy. Soc.* B, **104**, 206
Keilin, D. (1930) *Proc. roy. Soc.* B, **106**, 418
Keilin, D. & Hartree, E. F. (1937) *Proc. roy. Soc.* B, **122**, 298
Keilin, D. & Hartree, E. F. (1938) *Proc. roy. Soc.* B, **125**, 171
Keilin, D. & Hartree, E. F. (1939) *Proc. roy. Soc.* B, **127**, 167
Keilin, D. & Hartree, E. F. (1940) *Proc. roy. Soc.* B, **129**, 277
Keilin, D. & Hartree, E. F. (1949) *Biochem. J.* **44**, 205
Keilin, D. & Hartree, E. F. (1953) *Nature, Lond.* **171**, 413
Kubowitz, F. & Haas, E. (1932) *Biochem. Z.* **255**, 247
Lehninger, A. L. (1949) *J. biol. Chem.* **178**, 625
Lehninger, A. L. (1951) In: Edsall, J. T., ed. *Enzymes and enzyme systems: their state in nature.* Cumberlege, London
Mac Munn, C. A. (1884) *J. Physiol.* **5**, *Proc.* xxiv
Mac Munn, C. A. (1886) *Phil. Trans.* **177**, 267
Paul, K.-G. (1951) In: Sumner, J. B. & Myrbäck, K. D. R., ed. *The enzymes: chemistry and mechanism of action*, vol. 2, pt 1, p. 357. Academic Press, New York
Pflüger, E. (1875) *Pflüg. Arch. ges. Physiol.* **10**, 251
Potter, V. R. & Reif, A. E. (1952) *J. biol. Chem.* **194**, 287
Rawlinson, W. A. & Hale, J. H. (1949) *Biochem. J.* **45**, 247
Röhman, F. & Spitzer, W. (1895) *Ber. dtsch. chem. Ges.* **28**, 567
Slater, E. C. (1949a) *Biochem. J.* **44**, 305
Slater, E. C. (1949b) *Biochem. J.* **45**, 14
Slater, E. C. (1949c) *Biochem. J.* **45**, 8
Slater, E. C. (1950a) *Biochem. J.* **46**, 484
Slater, E. C. (1950b) *Nature, Lond.* **166**, 982
Spallanzani, L. (1807) In: Senebier, J., ed. *Rapports de l'air avec les êtres organisés*, vol. 2. Paschoud, Geneva

Continued at foot of page 97

SOME ASPECTS OF THE ENERGY TRANSFORMATION IN LIVING MATTER

H. A. KREBS M.D. M.A. F.R.S.

Medical Research Council Unit for Research in Cell Metabolism
Department of Biochemistry, University of Sheffield

1 The three phases of foodstuff degradation
2 The energy-yielding steps of intermediary metabolism
3 The build-up of phosphate bond energy
4 Energy transformations in different types of organism
5 Pathways of syntheses
6 " Natura enim simplex est "
References

The last decade has witnessed remarkable progress in the analysis of energy transformations in living organisms. A picture has emerged from the mass of detailed information, which reveals some striking characteristics of the chemical organization of living matter.

The concept is now firmly established (Lipmann, 1941) that the chemical energy set free in the breakdown of foodstuffs (unless it generates heat) is transformed into a special kind of chemical energy before it is converted into other forms of energy, such as mechanical work in muscle or osmotic work in secreting glands. This special form of chemical energy is that stored in the pyrophosphate bonds of adenosine-triphosphate (ATP).

Living matter shares the requirement of a special fuel with man-made machines. In a steam engine the energy obtained from the combustion of a variety of fuels is converted into steam pressure, before it can do mechanical work. In an internal combustion engine it is the high pressure and temperature of the gas resulting from the combustion of various fuels which carries out work. An electric motor must be supplied with electrical energy in a specified form to do work.

The first major stage, then, of the energy transformations in living matter culminates in the synthesis of pyrophosphate bonds of ATP, at the expense of the free energy of the degradation of foodstuffs. The over-all thermodynamic efficiency of this process is estimated at 60–70%. This is high when compared with the efficiency of man-made machines depending on the burning of a fuel as a source of energy. The greater efficiency is possible because living matter is not a heat engine, but a " chemical " engine which is organized in a very special manner.

1. The Three Phases of Foodstuff Degradation

The production of energy by the combustion of foodstuffs in living matter may be said to proceed in three major phases (Table I).

In phase I the large molecules of the food are broken down to the small constituent units. Proteins are converted to amino-acids, carbohydrates to hexoses and fats to glycerol and fatty acids. The amounts of energy liberated in the stages of phase I are relatively small. The free energy of hydrolysis of the glucosidic bonds of starch or glycogen is about 4.3 kg. cal. (Burton & Krebs, 1953), of the peptide bonds about 3.0 kg. cal. (Huffman, 1942), and of an ester about 2.5 kg. cal., per mole (see, for example, Kaplan, 1951). This means that about 0.6% of the free energy of polysaccharides and proteins, and about 0.1% of that of triglyceride fats, is liberated in phase I. These quantities are not utilized except for the generation of heat. The reactions of this phase thus do not yield utilizable energy: they merely prepare the foodstuffs for the energy-giving processes proper. The reactions of phase I take place in the intestinal tract, and also in tissues when reserve material is mobilized for energy production.

In phase II the diversity of small molecules produced in the first phase—three or more different hexoses, glycerol, about twenty amino-acids and a number of fatty acids—are incompletely burned, the end-product being, apart from carbon dioxide and water, one of three substances: acetic acid in the form of acetyl-coenzyme A, α-ketoglutarate, or oxaloacetate. The first of these three constitutes the greater amount: two-thirds of the carbon of carbohydrate and of glycerol, all carbon atoms of the common fatty acids and approximately half the carbon skeleton of amino-acids yield acetyl-coenzyme A. α-Ketoglutaric acid arises from glutamic acid, histidine, arginine, citrulline, ornithine and proline (see Krebs, 1953); oxaloacetic acid from aspartic acid, and through malic acid from part of the benzene ring of tyrosine and phenylalanine. The details of the pathways cannot be discussed here in full; they are outlined in Table I. One matter of principle, however, must be emphasized: the number of steps which living matter employs in order to reduce a great variety of different substances to three basic units is astonishingly small and could certainly not be equalled with the tools at present available to the organic chemist.

The three end-products of the second phase are metabolically closely interrelated. They take part in the phase III of energy production: the tricarboxylic acid cycle, the common " terminal " pathway of oxidation of all foodstuffs. The cycle is shown in Table II; it sets out how one acetic acid equivalent is burned to carbon dioxide and water, a series of di- and tri-carboxylic acids appearing as intermediate stages.

As each step of the intermediary metabolism requires a specific enzyme, it is evident that a common pathway of oxidation results in an economy of chemical tools. Surveying

---

CYTOCHROME

Continued from page 96

Spitzer, W. (1897) *Pflüg. Arch. ges. Physiol.* **67**, 615
Szent-Györgyi, A. (1924) *Biochem. Z.* **150**, 195
Tamiya, H. & Yamagutchi, S. (1933) *Acta phytochim., Tokyo*, **7**, 233
Theorell, H. & Åkesson, A. (1941) *J. Amer. chem. Soc.* **63**, 1804; 1812; 1818
Thunberg, T. (1937) *Ergebn. Physiol.* **39**, 76
Tsou, C. L. (1951) *Biochem. J.* **49**, 512
Warburg, O. (1913) *Pflüg. Arch. ges. Physiol.* **154**, 599
Warburg, O. (1925) *Ber. dtsch. chem. Ges.* **58**, 1001
Warburg, O. (1926) *Biochem. Z.* **177**, 471
Warburg, O. & Gewitz, H.-S. (1952) *Hoppe-Seyl. Z.* **288**, 1
Warburg, O. & Negelein, E. (1929) *Biochem. Z.* **214**, 64
Wieland, H. (1922) *Ergebn. Physiol.* **20**, 506
Yaoi, H. & Tamiya, H. (1928) *Proc. imp. Acad. Japan*, **4**, 436

TABLE I. THE THREE MAIN PHASES OF ENERGY PRODUCTION FROM FOODSTUFFS

Phase	Outline of chemical change
I	carbohydrates → hexoses; proteins → about 20 amino-acids; fats → glycerol; fatty acids
II	hexoses, glycerol → lactic acid → pyruvic acid; several amino-acids (alanine, serine, cysteine) → pyruvic acid; fatty acids → acetyl-coenzyme A; several amino-acids (3 leucines, tyrosine, phenylalanine) → acetyl-coenzyme A; several amino-acids (glutamic acid, histidine, prolines, arginine) → α-ketoglutaric acid; several amino-acids → oxaloacetic acid
III	acetyl-coenzyme A, α-ketoglutaric acid, oxaloacetic acid → tricarboxylic acid cycle

TABLE II. THE TRICARBOXYLIC ACID CYCLE

Substances which enter the cycle (acetyl-coenzyme A, water) are written inside, substances which arise are written outside the cycle. During one turn of the cycle, one acetic acid equivalent is completely oxidized. The four pairs of H atoms that arise react with O_2 to form water. (For further details see Krebs, 1950, 1953; Ochoa, Stern & Schneider, 1951; Stern *et al.* 1951.)

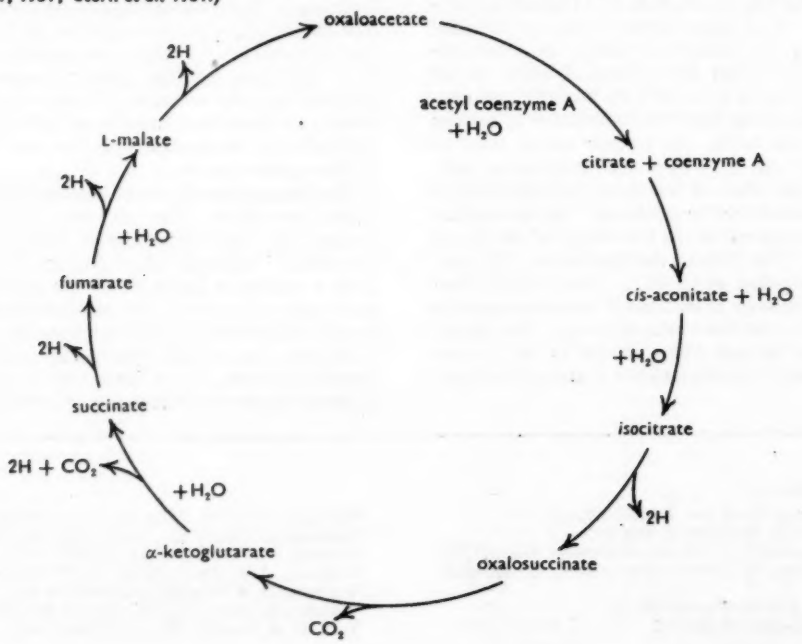
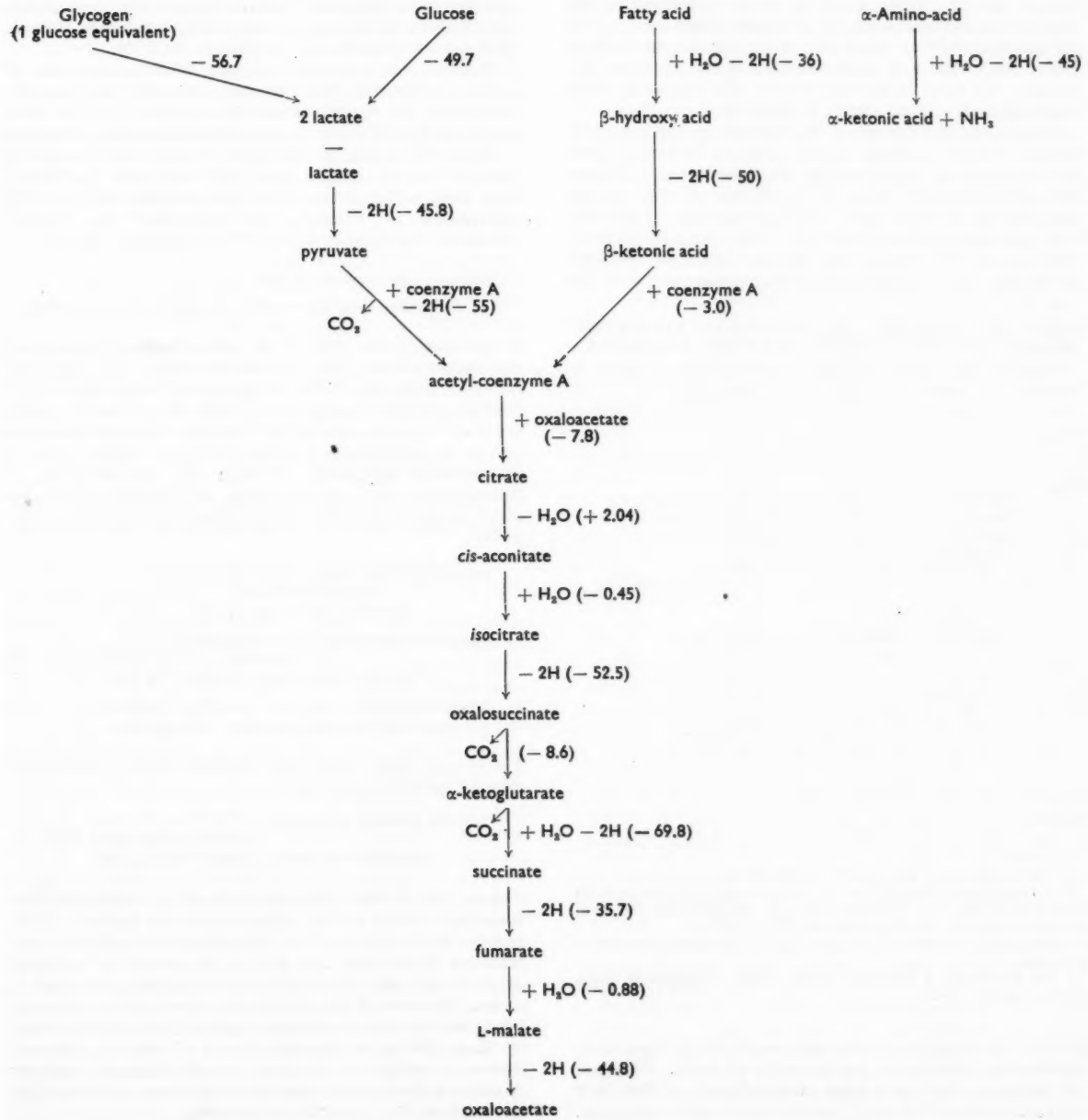


TABLE III. FREE ENERGY CHANGES IN THE BREAKDOWN OF FOODSTUFFS

The figures in brackets refer to the free energy change ΔG in kilogram-calories per mole reactants at 25° C.; pH 7.0; 0.2 atm. O_2; 0.05 atm. CO_2; concentrations of other reactants at 0.01 M activity. The pairs of H atoms removed are assumed to be oxidized to H_2O by molecular oxygen. Fatty acids are probably oxidized as acyl-coenzyme A derivatives. The assumption that β-hydroxy acids are intermediates in the oxidation of fatty acids is arbitrary, but the free energy change will not be very different if instead unsaturated acids are the intermediates. The value given for the oxidative deamination of amino-acids is an average figure varying a little from amino-acid to amino-acid. (For further particulars see Burton & Krebs, 1953; Kaplan, 1951.)



the pathway of the degradation of foodstuffs as a whole, one cannot but be impressed by the relative simplicity of the arrangement. The total number of steps required to release the available energy from a multitude of different substrates is unexpectedly small.

2. The Energy-Yielding Steps of Intermediary Metabolism

But the economy of stages does not end here. Energy is not released at every intermediary step of metabolism. Most, though not all, of the available energy contained in the foodstuffs is liberated when the hydrogen atoms removed by the dehydrogenations react with molecular oxygen to form water. Such reactions occur in phase II and in phase III. Roughly one-third of the total energy of combustion is set free in phase II and two-thirds in phase III.

Data on the magnitude of the free energy changes have recently become available and are given in Table III. Even with a common terminal pathway of oxidation and the shortness of the routes in phase II, the number of stages where utilizable energy is set free is still considerable. There are four such stages in the tricarboxylic acid cycle. In phase II there are at least twenty-nine different oxidative reactions (Table IV). Their number cannot be precisely stated as the pathway of oxidation of some amino-acids (e.g., tryptophan, methionine, lysine) is not yet known in all detail. There are two oxidative stages each when carbohydrates or when fatty acids are prepared for entry into the tricarboxylic acid cycle.

TABLE IV. NUMBER OF DIFFERENT OXIDATIVE STAGES AT WHICH ENERGY BECOMES AVAILABLE

"Phase" refers to the classification given in Table I; energy is liberated when molecular oxygen is the oxidizing agent.

Phase I	0	
Phase II		
carbohydrate ⟶ acetyl-coenzyme A	2	29
fatty acids ⟶ acetyl-coenzyme A	2	
glycerol* ⟶ acetyl-coenzyme A	1	
amino-acids		
oxidative deamination of amino-acids†	(about) 16	
oxidative decarboxylation of corresponding α-ketonic acids‡	0	
additional oxidation of phenylalanine and tyrosine leading to malic and acetoacetic acids§	4	
oxidation of proline and hydroxyproline to glutamic acid	2	
oxidation of arginine to glutamic acid	2	
conversion of histidine to glutamic acid	0	
conversion of carbon chain of leucines to acetyl-coenzyme A	0	
Phase III	4	

* Two more stages shared with carbohydrate.

† A figure lower than the number of amino-acids is given because some amino-acids may share one step, e.g., glutamic acid and histidine; serine, cysteine and glycine (see Bach, 1952).

‡ Assuming mechanism to be same as the analogous reaction of pyruvate.

§ See Weinhouse & Millington (1948, 1949); Schepartz & Gurin (1949); Lerner (1949).

One special stage is required for glycerol, two others being identical with those of carbohydrate. Amino-acids, with a few exceptions, require at least one oxidative step each—the oxidative deamination. In the cases of glutamic acid, aspartic acid, and alanine, this one step is sufficient to produce an intermediate also encountered in either carbohydrate breakdown or in the tricarboxylic acid cycle. The exceptional amino-acids referred to above, which do not require special oxidative steps, include histidine, which yields glutamic acid anaerobically, and cysteine and glycine, which are both convertible into serine. In many other cases additional steps are required. Details about the number of energy-releasing steps in phase II of amino-acid metabolism are given in Table IV.

There are thus some thirty stages at which major parcels of energy are liberated. This is already a relatively small number, considering the variety of starting materials, but by some simple devices the number receives a further drastic reduction.

Most of the stages listed in Table IV are complex reactions involving several distinct steps. In many cases the first of these steps is the interaction of the substrates with pyridine nucleotides, as Warburg and colleagues first showed (Warburg, Christian & Griese, 1935; Warburg, 1938):

$$\text{substrate} + \text{pyridine nucleotide} \rightleftharpoons \text{dehydrogenated substrate} + \text{reduced pyridine nucleotide}$$

It follows from the data on the oxidoreduction potentials of the reactants (Ball, 1944; Burton & Wilson, 1952, 1953) that no appreciable amounts of energy are set free in this type of reaction, except when α-ketonic acids are oxidized, a case discussed below in more detail. The free energy of the oxidation of the substrates is liberated when the reduced forms of the pyridine nucleotides, through the intermediation of flavoproteins and iron porphyrins, are reoxidized by molecular oxygen, the main stages being the following (Ball, 1944):

$$\text{reduced pyridine nucleotide} + \text{flavoprotein} \longrightarrow \text{reduced flavoprotein} + \text{pyridine nucleotide} \quad (1)$$
(standard free energy change, −11 kg. cal.)

$$\text{reduced flavoprotein} + 2\ \text{ferricytochrome} \longrightarrow \text{flavoprotein} + 2\ \text{ferrocytochrome} \quad (2)$$
(standard free energy change, −16 kg. cal.)

$$2\ \text{ferrocytochrome} + \tfrac{1}{2}\ O_2 \longrightarrow 2\ \text{ferricytochrome} \quad (3)$$
(standard free energy change, −25 kg. cal.)

The over-all effect of the three reactions is the oxidation of reduced pyridine nucleotide:

$$\text{reduced pyridine nucleotide} + \tfrac{1}{2}\ O_2 \longrightarrow \text{pyridine nucleotide} + H_2O \quad (4)$$
(standard free energy change, −52 kg. cal.)

This scheme of three reactions, it should be emphasized, is a simplified version of the actual events (see Herbert, 1950), because more than one iron porphyrin and probably more than one flavoprotein take part in the transfer of hydrogen or electrons, these intermediate carriers being arranged in series. The chain of the catalysts may also include substances other than the three main types, such as the factor discovered by Slater (1949), the chemical identity of which is unknown. Moreover, as there are two pyridine nucleotides and a number of different flavoproteins and iron porphyrins, several variants of reactions (1), (2) and (3) are possible.

The stages (1), (2) and (3) are common to many of the oxidations in phases II and III: to three out of the four oxidative steps of the tricarboxylic acid cycle; to the oxidation of lactate and glycerol to pyruvate; to the oxidation of α-ketonic acids to fatty acids and carbon dioxide; and probably also to the stages involved in the β-oxidation of fatty acids. Information on the enzymic mechanism of β-oxidation is still incomplete, but the available evidence suggests that diphosphopyridine nucleotide (DPN) is involved (Burton, 1948; Kennedy & Lehninger, 1951). One stage of the four dehydrogenations of the tricarboxylic acid cycle, viz., the dehydrogenation of succinic acid, does not require a pyridine nucleotide or flavoprotein, but is catalysed by iron porphyrins only. In this case, only the last step of the series (1), (2) and (3) is shared with the other dehydrogenase systems. This may be conditioned by the thermodynamic characteristics of the succinic dehydrogenase system, whose oxidoreduction potential is less negative than that of any other substrate dehydrogenase, with the possible exception of the fatty acid dehydrogenase system, which is expected to be of the same order as that of succinic dehydrogenase.

Thus, with one major and possibly a few minor exceptions, all stages listed in Table IV release energy through the same series of reactions. The major exception is the deamination of amino-acids, but this exception is probably only an apparent one. One amino-acid—glutamic acid—can actually interact with pyridine nucleotides, and virtually all amino-acids, according to a discovery made by Braunstein (1947; see also Cammarata & Cohen, 1950), can transfer their amino-group to α-ketoglutarate by "transamination", a reversible exchange of the keto and amino groups:

$$\text{amino-acid} + \alpha\text{-ketoglutarate} \rightleftharpoons \text{keto-acid} + \text{glutamate} \quad (5)$$

This reaction links the deamination of all amino-acids with the dehydrogenation of glutamic acid by pyridine nucleotides:

$$\text{glutamate} + \text{pyridine nucleotide} \rightleftharpoons \alpha\text{-ketoglutarate} + NH_3 + \text{reduced pyridine nucleotide} \quad (6)$$

Braunstein has suggested that reactions (5) and (6) represent the main mechanism of deamination, but this still remains to be proved. It is an attractive hypothesis in connexion with the thesis put forward in this paper. If it holds, it would mean that no free energy is released in the transamination reaction (5), as this reaction is readily reversible and the concentration of the reactants remains virtually constant. Nor is there a free energy change when reaction (6) occurs under physiological conditions. Thus the free energy of the oxidative deamination of amino-acids generally would be released when the reduced pyridine nucleotide undergoes oxidation through reactions (1), (2) and (3), and the 16 energy-yielding reactions of amino-acids listed in Table IV would be eliminated as separate processes. It is another attractive aspect of this concept that it ascribes a biological significance to the transaminases—enzymes the metabolic role of which is otherwise not understood.

Whatever the merits of Braunstein's hypothesis, the picture does not fundamentally change even if there is an alternative mechanism of oxidative deamination, because this would still involve flavoproteins and iron porphyrins. Amino-acid oxidases not requiring pyridine nucleotides have in fact been isolated from animal tissues, but the activity of those amino-acid oxidases which attack the more common amino-acids in animal tissues is weak and the physiological significance of the enzymes is not clearly established (see Krebs, 1951). In the kidney they may contribute to the formation of ammonia required for excretion. The amino-acid oxidases are flavoproteins, linked in situ with molecular oxygen by iron porphyrins, as indicated by their sensitivity to cyanide. Hence, the last two of the stages (1), (2) and (3) still represent the steps where energy is released when amino-acids are oxidized. In most tissues the activities of the transaminases and glutamic dehydrogenase are very much greater than those of the amino-acid oxidases, and some tissues, like cardiac muscle, are even devoid of amino-acid oxidases and yet utilize amino-acids (Clarke & Whaler, 1952). These facts suggest that Braunstein's mechanism is a major pathway of amino-acid oxidation.

There are a few oxidative steps in the breakdown of amino-acids, for example, of tyrosine, phenylalanine, the prolines, methionine and the ω-deamination of lysine and ornithine, where the enzymic mechanisms are not yet known. It is possible that pyridine nucleotides, flavoproteins and iron porphyrins are the catalytic agents, in which case these reactions would comply with the general pattern. But it is also feasible that the free energy of these specialized reactions is not utilized and "lost" as heat. This would account for part of the rise of the basal metabolic rate when protein serves as the source of energy—the "specific dynamic action of protein". As the amino-acids in question form a relatively small proportion of the protein molecule the loss of energy would be small.

To sum up, the energy-yielding oxidative stages of Table IV are all complex reactions, which involve a specific primary step, or series of steps—the reaction with pyridine nucleotide and, in the case of amino-acids, transamination—and a set of intermediary hydrogen carriers, reacting according to (1), (2) and (3), common to all major stages. The release of energy occurs along this common pathway and not when the substrate reacts. This arrangement reduces the total number of energy-yielding steps to seven, or more correctly to seven types.

3. The Build-Up of Phosphate Bond Energy

Of the seven energy-yielding reactions proper, three occur when reduced pyridine nucleotide is oxidized by molecular oxygen through the stages (1), (2) and (3). A fourth is the oxidation of succinate by cytochrome *c*. Two more occur in anaerobic glycolysis and the seventh during the oxidative decarboxylation of α-ketonic acids.

The crucial feature of these reactions, from the biological point of view, is not the liberation of energy, but the coupling of energy production with energy utilization, i.e., with the synthesis of pyrophosphate bonds of adenosinetriphosphate. The utilization of the free energy of a chemical process requires a coupling mechanism specifically adjusted to the circumstances of the energy-yielding steps. By the reduction of the number of such steps to a minimum, living matter reduces the requirement for specific mechanisms and thereby effects a striking economy.

The mechanism of energy transfer is known for two of the seven energy-giving reactions—those of anaerobic glycolysis. For another, the oxidative decarboxylation of α-ketonic acids, it is known in principle, but some uncertainty of detail remains. It is completely unknown for the remaining four reactions.

Following a suggestion of Lipmann (1941, 1946), the pyrophosphate bonds of adenosinetriphosphate are commonly referred to as " energy-rich " bonds. This is a convenient term from the biological point of view (although it does not tally with the use of the term " bond energy " in physical chemistry, where it refers to the energy required to disrupt a bond between two atoms). It expresses the fact that the hydrolysis of this bond releases a much larger amount of free energy, viz., 10–15 kg. cal. per mole, than the usual hydrolytic reactions occurring in living matter, which release 2–4 kg. cal. per mole. The energy stored in pyrophosphate bonds is often referred to, again according to Lipmann, as " phosphate bond energy ".

The generation of phosphate bond energy in the three known examples is shown in Tables V, VI and VII. The mechanism in all three cases follows the same pattern. A series of reactions is so conducted as to convert an ordinary " low energy bond ", such as the ester bond of 2-phosphoglyceric acid or the C—S bond of the adducts of mercaptans and aldehydes or ketones, into an energy-rich bond. The first energy-rich bond to arise is not a pyrophosphate bond, but an acyl-mercaptan bond in Case I (Table V) and Case III (Table VII) and an *enol* phosphate in Case II (Table VI). By a reversible transfer reaction the acyl-mercaptans are converted into other energy-rich bonds, either acyl phosphate (Case I) or a phosphoryl-mercaptan (Case III). By further transfer reactions phosphate is conveyed to adenosinediphosphate forming ATP. With the exception of the *enol* phosphate formed in Case II the various energy-rich bonds are all acid anhydrides. In Case I the energy-rich bond arises by the dehydrogenation of the addition compound of substrate and the active group of the enzyme, an SH group of glutathione. In Case II it is the elimination of the elements of water from 2-phosphoglyceric acid which creates an energy-rich bond. In Case III (see Table VII) the reaction generating an energy-rich bond appears to have many features in common with the first. Again an SH compound—in this case coenzyme A—forms an addition compound with the substrate and undergoes dehydrogenation, thereby producing an energy-rich acyl-mercaptan. As in Case I, this undergoes phosphorolysis; the result in this case is not a carboxyl phosphate, but an energy-rich phosphoryl-mercaptan. This scheme is to be looked upon as a tentative hypothesis.

The three reactions just discussed all occur anaerobically and involve substrate molecules, and are therefore often referred to as " anaerobic phosphorylations at substrate level ". In contrast, the phosphorylation associated with the oxidation of reduced pyridine nucleotide does not directly involve substrates but intermediary hydrogen and electron carriers; it is usually observed when the carriers are present in catalytic quantities which are continuously reduced by the substrate and reoxidized by molecular oxygen. Because of the involvement of oxygen the mechanism is commonly referred to as " oxidative phosphorylation ".

The characteristics of oxidative phosphorylation differ in important aspects from those of the anaerobic phosphorylation at substrate level. Unlike the latter, oxidative phosphorylation is inhibited by nitrophenols, halogen phenols, azide, gramicidin and other substances. At suitable concentrations these inhibitors of phosphorylation do not inhibit the oxidations; they thus " uncouple " oxidation and phosphorylation.

TABLE V. FORMATION OF ENERGY-RICH PHOSPHATE BONDS IN ANAEROBIC GLYCOLYSIS : CASE I

According to Warburg & Christian (1939), Bücher (1947) and Racker & Krimsky (1952), energy-rich phosphate is formed during the conversion of glyceraldehyde-3-phosphate into glyceric acid-3-phosphate.

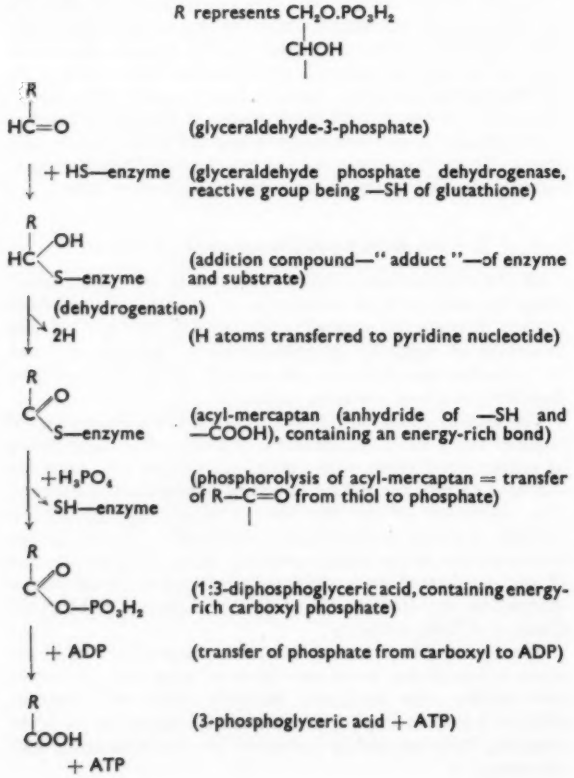


TABLE VI. FORMATION OF ENERGY-RICH PHOSPHATE BONDS IN ANAEROBIC GLYCOLYSIS : CASE II

Energy-rich phosphate is formed during conversion of 2-phosphoglyceric acid into phosphoe*nol*pyruvic acid.

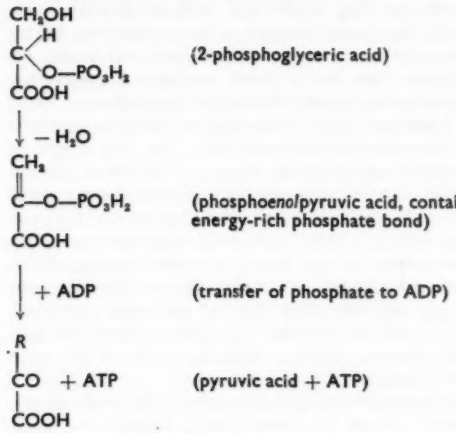
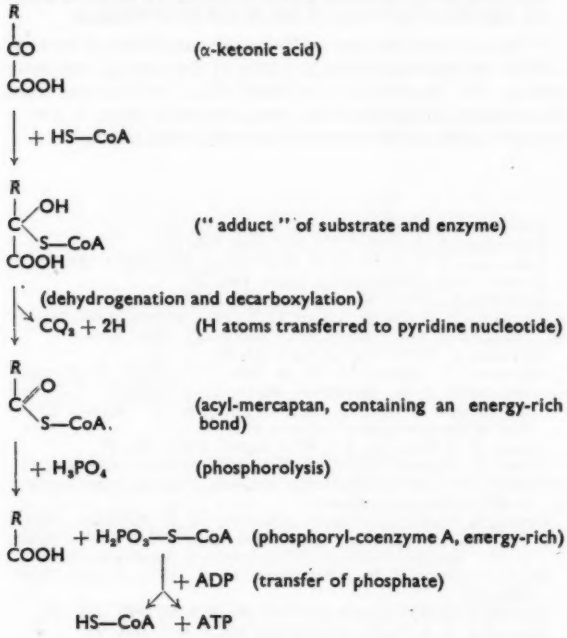


TABLE VII. FORMATION OF ENERGY-RICH PHOSPHATE BONDS DURING THE OXIDATIVE DECARBOXYLATION OF α-KETONIC ACID : CASE III

Tentative scheme according to Lynen & Reichert (1951), Lynen, Reichert & Rueff (1951), Lipmann, Jones & Black (1952), Green (1952), Ochoa (1952). "HS—CoA" stands for coenzyme A acting as the prosthetic group of the enzyme.



No comparable effect is known in anaerobic phosphorylation. Furthermore, whilst the anaerobic phosphorylation occurs in homogeneous solutions, the mechanism responsible for oxidative phosphorylation is associated with insoluble particles of the cytoplasm, the mitochondria. It cannot be brought into solution and it appears that it functions only in a heterogeneous system. The mechanisms of oxidative and anaerobic phosphorylation thus seem to differ in essential points.

The standard free energy change of reaction (4) is about 52 kg. cal., which is sufficient for the synthesis of at least three pyrophosphate bonds each requiring the addition of about 12 to 14 kg. cal. under physiological conditions. Measurements of the ratio

$$\frac{\text{equivalents of organic phosphate formed}}{\text{atoms of oxygen consumed}}$$

show that three pyrophosphate bonds can actually be synthesized when one molecule of reduced pyridine nucleotide is oxidized (Lehninger, 1949, 1951; Lehninger & Smith, 1949) and it is considered as probable, though not as certain (see Slater, 1950), that one each is formed for each of the three component reactions (1), (2) and (3).

4. Energy Transformations in Different Types of Organism

The same basic mechanism of energy production has been found in all animal species from protozoa to mammals. It has also been demonstrated in micro-organisms and plants. Alternative mechanisms may exist, especially in micro-organisms, but these alternatives are generally relatively minor modifications of the same general theme. Thus the tricarboxylic acid cycle is probably replaced in certain micro-organisms by other pathways (see Krebs, Eggleston & Gurin, 1952); but where this is the case the role of pyridine nucleotides, flavoproteins and iron porphyrins is still the same. Some micro-organisms, such as *Pseudomonas* species, are capable of deriving energy from a large variety of substrates, including aromatic and other substances not utilizable by the majority of living organisms. There is every reason for believing that the actual process giving energy in *Pseudomonas* is the same as in other types of organisms, viz., reaction (4). The special feature of the organisms is the ability to produce dehydrogenases capable of transferring hydrogen to pyridine nucleotide from unusual substrates.

It may be inferred from these considerations that the basic mechanism of energy utilization must have arisen at a very early stage of the evolution of life.

5. Pathways of Syntheses

The primary reactions by which the foodstuffs are degraded can all be readily reversed, some directly because the reactants are in thermodynamic equilibrium, others through a cyclic process like the tricarboxylic acid cycle where the pathways of degradation and synthesis are different. "Primary" reactions here refer to the reactions involving the substrate, i.e., the organic molecules derived from foodstuffs, but do not refer to the interaction between molecular oxygen and the iron porphyrins of the cytochrome system, which reaction, for all practical purposes, is irreversible in non-photosynthetic organisms.

The reversibility at substrate level means that the very routes which serve as pathways of degradation may also lead to syntheses. Acetyl-coenzyme A is not only the chief common breakdown product but, together with the intermediates of the tricarboxylic acid cycle, the basic unit for organic syntheses. The carbon skeletons of the fatty acids, cholesterol, steroid hormones, amino-acids and porphyrins, can all be synthesized, or partially synthesized, from acetyl-coenzyme A and the intermediates of the tricarboxylic acid cycle (see Lipmann, 1950). The simplicity of the synthetic mechanisms thus mirrors the simplicity of the degradative mechanisms in living matter.

6. "Natura enim simplex est"

This survey is confined to a discussion of the first stage of energy transformations in living matter, in which a multitude of chemical fuels is converted into one type of chemical fuel. It is not intended to consider here the second stage in which phosphate bond energy is converted into other forms of energy.

The work surveyed in this article has revealed a measure of simplicity in the arrangements for energy transformations that was hardly to be expected. By economy in the number of steps of degradation and by employment of the pyridine nucleotides as a common primary oxidizing agent (which reacts with the substrates without loss of energy), the total number of processes giving energy, and therefore of mechanisms transforming energy, is reduced to the low figure of seven

UNIVERSITY OF MICHIGAN LIBRARIES

types. An economy of tools, and thereby of space, may be looked upon as an essential feature of the make-up of living matter.

The kind of simplicity here encountered is in accord with evolutionary theory and is repeated at many other levels of the chemical organization of living matter. The amazing variety of living forms which endows every individual of the higher organisms with its own distinguishable characteristics tends to obscure the fact that relatively simple principles may hide behind complex phenomena. Virtually infinite variety can be the result of the permutation and combination of a relatively small number of different basic units. The individuality of organisms is essentially an individuality of proteins and all proteins are but permutations and combinations of some twenty constituent amino-acids, all of which are derivatives of one parent substance—glycine. As Kossel (1912) has pointed out some forty years ago:

> We obtain some idea of the possible variety in the combination and reactions of the protein constituents by recalling the fact that they are as numerous as the letters of the alphabet which are capable of expressing an infinite number of thoughts.

" Natura enim simplex est " was the conviction of Newton (1687), on which he based his study of the laws of inanimate nature. It is becoming increasingly clear, in particular from the progress of biochemistry, that Newton's thesis is also a valuable guide in the study of the phenomena of life.

References

Bach, S. J. (1952) *The metabolism of protein constituents in the mammalian body*. Cumberlege, London

Ball, E. G. (1944) *Ann. N.Y. Acad. Sci.* **45,** 363

Braunstein, A. E. (1947) *Advanc. Protein Chem.* **3,** 1

Bücher, T. (1947) *Biochim. biophys. Acta,* **1,** 292

Burton, K. (1948) *Nature, Lond.* **161,** 606

Burton, K. & Krebs, H. A. (1953) *Biochem. J.* **54,** 94

Burton, K. & Wilson, T. H. (1952) *Biochim. biophys. Acta,* **8,** 114

Burton, K. & Wilson, T. H. (1953) *Biochem. J.* **54,** 86

Cammarata, P. S. & Cohen, P. P. (1950) *J. biol. Chem.* **187,** 439

Clarke, E. W. & Whaler, B. C. (1952) *J. Physiol.* **117,** 9*P*

Green, D. E. (1952) *II^e Congrès International de Biochimie, Paris, 21–27 Juillet 1952.* In: *Symposium sur le cycle tricarboxylique,* p. 5. Société d'Edition d'Enseignement Supérieur, Paris

Herbert, D. (1950) *Rep. Progr. Chem.* **47,** 335

Huffman, H. M. (1942) *J. phys. Chem.* **46,** 886

Kaplan, O. (1951) In: Sumner, J. B. & Myrbäck, K. D. R., ed. *The enzymes: chemistry and mechanism of action,* vol. 2, pt 1, p. 55. Academic Press, New York

Kennedy, E. P. & Lehninger, A. L. (1951) *J. biol. Chem.* **190,** 361

Kossel, A. (1912) *Johns Hopk. Hosp. Bull.* **23,** 65. Quoted by Bach, S. J. (1952)

Krebs, H. A. (1950) *Harvey Lect.* **44,** 166

Krebs, H. A. (1951) In: Sumner, J. B. & Myrbäck, K. D. R., ed. *The enzymes: chemistry and mechanism of action,* vol. 2, pt 1, p. 499. Academic Press, New York

Krebs, H. A. (1953) In: Greenberg, D. M., ed. *The tricarboxylic acid cycle, chemical pathways of metabolism.* Academic Press, New York

Krebs, H. A., Eggleston, L. V. & Gurin, S. (1952) *Biochem. J.* **51,** 614

Lehninger, A. L. (1949) *J. biol. Chem.* **178,** 625

Lehninger, A. L. (1951) *J. biol. Chem.* **190,** 345

Lehninger, A. L. & Smith, S. W. (1949) *J. biol. Chem.* **181,** 415

Lerner, A. B. (1949) *J. biol. Chem.* **181,** 281

Lipmann, F. (1941) *Advanc. Enzymol.* **1,** 99

Lipmann, F. (1946) In: Green, D. E., ed. *Currents in biochemica research,* p. 137. Interscience Publishers, New York

Lipmann, F. (1950) *Harvey Lect.* **44,** 99

Lipmann, F., Jones, M. E. & Black, S. (1952) *II^e Congrès International de Biochimie, Paris, 21–27 Juillet 1952.* In: *Symposium sur le cycle tricarboxylique,* p. 55. Société d'Edition d'Enseignement Supérieur, Paris

Lynen, F. & Reichert, E. (1951) *Angew. Chem.* **63,** 47

Lynen, F., Reichert, E. & Rueff, L. (1951) *Liebigs Ann.* **574,** 1

Newton, I. (1687) *Philosophiae naturalis principia mathematica,* p. 402. Streater, London

Ochoa, S. (1952) *II^e Congrès International de Biochimie, Paris, 21–27 Juillet 1952.* In: *Symposium sur le cycle tricarboxylique,* p. 73. Société d'Edition d'Enseignement Supérieur, Paris

Ochoa, S., Stern, J. R. & Schneider, M. C. (1951) *J. biol. Chem.* **193,** 691

Racker, E. & Krimsky, I. (1952) *J. biol. Chem.* **198,** 721; 731

Schepartz, B. & Gurin, S. (1949) *J. biol. Chem.* **180,** 663

Slater, E. C. (1949) *Biochem. J.* **45,** 1; 8; 14

Slater, E. C. (1950) *Nature, Lond.* **166,** 982

Stern, J. R., Shapiro, B., Stadtman, E. R. & Ochoa, S. (1951) *J. biol. Chem.* **193,** 703

Warburg, O. (1938) *Ergebn. Enzymforsch.* **1,** 210

Warburg, O. & Christian, W. (1939) *Biochem. Z.* **303,** 46

Warburg, O., Christian, W. & Griese, A. (1935) *Biochem. Z.* **282,** 157

Weinhouse, S. & Millington, R. H. (1948) *J. biol. Chem.* **175,** 995

Weinhouse, S. & Millington, R. H. (1949) *J. biol. Chem.* **181,** 645

* * *

ALTERNATIVE ROUTES OF CARBOHYDRATE OXIDATION

FRANK DICKENS M.A. D.Sc. F.R.S.

Philip Hill Professor of Experimental Biochemistry
Middlesex Hospital Medical School, London

1 The basic pattern: anaerobic breakdown of carbohydrate
2 Oxidation of pyruvate formed from carbohydrate
3 Direct oxidation of hexosemonophosphate
4 Oxidation of 6-phosphogluconate: formation of pentose from hexose
5 Further metabolism of pentosephosphate
6 Part played by a C_7 sugar in the anaerobic transformation of pentose to hexose
7 Distribution of these metabolic pathways in nature and their relative significance
8 Summary
References

Knowledge of mechanism of carbohydrate metabolism was for a long time limited to fermentative breakdown not involving the interaction of atmospheric oxygen, and this forms the foundation of all later work on carbohydrate oxidation.

1. The Basic Pattern : Anaerobic Breakdown of Carbohydrate

The essential part played in alcoholic fermentation by phosphorylated sugars was first shown at the Lister Institute of Preventive Medicine, London, by Harden & Young (1909), who described the function of fructose-1:6-diphosphate (the Harden-Young ester), and by Robison (1922), who showed that hexosemonophosphates also participate in the form of the equilibrium mixture (Robison ester) of glucose-6-phosphate and fructose-6-phosphate. Harden and Young also described for the first time the coenzyme of fermentation, which they named cozymase, the structure of which was later elucidated by von Euler and Warburg (for a review see Schlenk, 1951) as diphospho-nicotinamide-adenine dinucleotide (DPN).

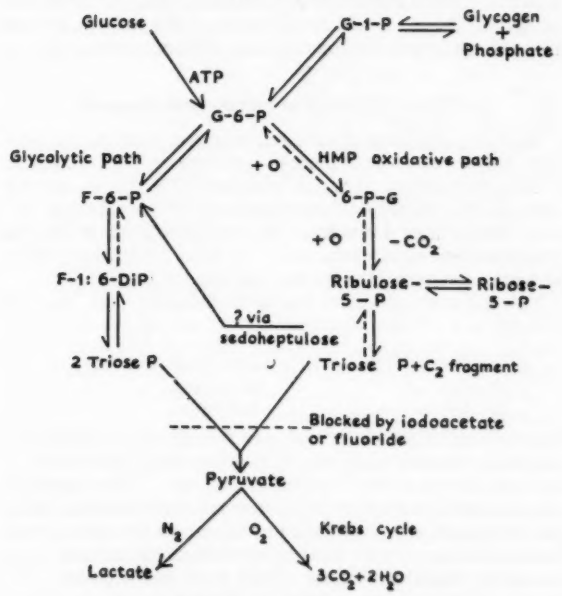
Meanwhile, in muscle metabolism similar phosphorylated sugar intermediates were found to be concerned, and the strikingly close basic patterns of energy metabolism in a very wide range of biological species came slowly to be recognized. Thus yeast metabolism was found to differ from that of muscle primarily in the presence in yeast of an active enzyme system (consisting of carboxylase and its coenzyme cocarboxylase) which under anaerobic conditions decarboxylated pyruvate to acetaldehyde, the source of the alcohol formed. In animal tissues, however, pyruvate was reduced directly by the lactic dehydrogenase-DPN system to lactic acid.

This anaerobic route of carbohydrate breakdown was at last, as the result of the work of Embden, Meyerhof, D. M. Needham, C. F. & G. T. Cori, and others, elucidated in detail and is generally known as the *glycolytic pathway* of carbohydrate degradation. The main reactions of this pathway are shown in fig. 1. The glycolytic system specifically requires cozymase (DPN) at the stages of oxidation by glyceraldehyde phosphate dehydrogenase and reduction by lactic dehydrogenase, and these two oxido-reductions balance one another in anaerobic glycolysis; this explains why the participation of atmospheric oxygen is not required. The glycolytic pathway is selectively poisoned at two stages, first by low concentrations of sulphydryl reagents, such as iodoacetate, which block the action of glyceraldehyde phosphate dehydrogenase, and second by fluoride ions, which block the action of the enzyme enolase (see fig. 1).

2. Oxidation of Pyruvate Formed from Carbohydrate

The pioneer work of R. A. Peters and his colleagues[1] at Oxford, extending over many years, gave the first insight into the role of vitamin B_1 (thiamine, aneurin) in *oxidation* of carbohydrate, particularly in brain tissue, the use of which greatly simplifies the problem since it has an almost exclusively carbohydrate metabolism and only very limited ability to utilize non-carbohydrate foodstuffs. As is well known, Peters (1936) found that the deficient respiration in brain tissue from B_1-avitaminous animals was restored by addition of thiamine and at the same time the pyruvate accumulation, due

FIG. 1. SOME ALTERNATIVE ROUTES OF CARBOHYDRATE METABOLISM : THE GLYCOLYTIC AND HEXOSEMONOPHOSPHATE OXIDATIVE ROUTES



ATP : adenosinetriphosphate
G-1-P : glucose-1-phosphate
G-6-P : glucose-6-phosphate
F-6-P : fructose-6-phosphate
F-1 : 6-DiP : fructose-1 : 6-diphosphate
HMP : hexosemonophosphate
6-P-G : 6-phosphogluconic acid

[1] See Peters, p. 116 of this Bulletin.

UNIVERSITY OF MICHIGAN LIBRARIES

to defective metabolism of the keto acid, was relieved. Later the function of thiamine pyrophosphate as the coenzyme of carboxylase was recognized, and it became clear that the carboxylase system was involved in further oxidation of pyruvate formed as an intermediate common to both the anaerobic and aerobic carbohydrate metabolism.

The detailed mechanism of pyruvate oxidation was then greatly clarified by the work of Krebs and his colleagues at Sheffield, who demonstrated in several tissues the part played by a series of enzymic reactions which followed the primary condensation of an acetate residue, derived from the oxidative decarboxylation of pyruvate, with oxaloacetic acid to form citric acid (see Krebs, 1943). The further oxidation of the citric acid thus formed involved transformation to *iso*citric acidand dehydroge nation of the latter, giving after decarboxylation ketoglutari cacid. Other enzyme systems oxidatively decarboxylate this acid to yield again oxaloacetate. Hence in each round of the Krebs cycle, one two-carbon unit derived from the oxidative decarboxylation of pyruvate is completely oxidized by a cyclical catalytic process, involving a number of steps which cannot be considered in detail here. Some of these require coenzyme A, others DPN as coenzyme, but in this oxidative cycle another coenzyme, containing a third phosphoric acid residue and known as triphospho-nicotinamide-adenine dinucleotide (TPN), is also required for the enzymic dehydrogenation of *iso*citric acid. This coenzyme was discovered by Warburg in his studies on the oxidation of hexosemonophosphate by extracts of red cells and of yeast; this work also introduced the possibility of an entirely different path of carbohydrate oxidation from that described in outline above. It will be necessary at this stage to discuss this alternative route of hexosemonophosphate oxidation.

3. Direct Oxidation of Hexosemonophosphate

Warburg, Christian & Griese (1935) described the preparation from yeast of a glucose-6-phosphate dehydrogenase (" Zwischenferment ") which required TPN for its activity and oxidized the hexosemonophosphate to 6-phosphogluconic acid, which they identified. The primary product of this oxidation has recently been shown by Cori & Lipmann (1952) to be the corresponding lactone and not the free acid.

This reaction proceeds stoicheiometrically, and may be written:

$$\underset{\text{(glucose-6-phosphate)}}{R.CHO} + TPN + H_2O = \underset{\text{(6-phospho-gluconate)}}{R.COOH} + TPNH_2$$

The reduced TPN so formed was reoxidized in Warburg's studies by the use of the flavoprotein oxidase discovered by him and known as the " old yellow enzyme ". The enzyme is not poisoned by cyanide, but in the cell the oxidation, being cyanide-sensitive, is presumed to proceed via the cytochrome system (Dickens, 1936), and in liver extracts has recently been shown by Glock & McLean (1952) to be mediated by cytochrome *c*-TPN reductase. This flavoprotein is situated mainly in the microsomes and mitochondria, and catalyses the reduction of cytochrome by reduced TPN. Liver is a rich source of glucose-6-phosphate dehydrogenase (Dickens & Glock, 1950, 1951), which occurs in the soluble enzyme fraction of the cytoplasm. Other sources of this enzyme (or of the 6-phosphogluconate dehydrogenase, described below, which usually appears to accompany the glucose-6-phosphate dehydrogenase in nature) besides liver, erythrocytes and yeast, include kidney, bone-marrow, spleen, heart, brain and liver carcinoma (Dickens, 1938a; Dickens & Glock, 1950, 1951; Horecker, 1951; Seegmiller & Horecker, 1951), as well as plant tissues and micro-organisms (Scott & Cohen, 1951; de Ley & Cornut, 1951; Gibbs, 1952; Horecker, 1953 and personal communication). The enzymes are thus widely distributed; they are, however, notably deficient in extracts of skeletal muscle. Since muscle tissue has been much used in the study of carbohydrate metabolism, this possibly accounts for the small attention hitherto given to the alternative pathway in animal metabolism.

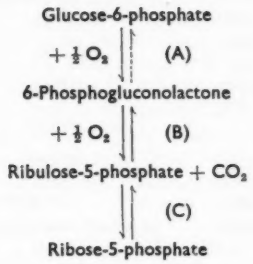
Horecker (1953) has recently demonstrated the reversal of this reaction, using the δ-lactone of 6-phosphogluconic acid (obtained by bromine oxidation of glucose-6-phosphate) in presence of reduced TPN, which was maintained in the reduced state by coupling it with another TPN enzymic system (e.g., *iso*citric dehydrogenase). One-third of the lactone could be reduced to glucose-6-phosphate under these conditions. It is quite probable that, in photosynthetic reactions where the necessary supply of reduced TPN could be sustained, this reversal could be of importance as a pathway to hexose.

4. Oxidation of 6-Phosphogluconate: Formation of Pentose from Hexose

The TPN-linked oxidation of glucose-6-phosphate can be caused to proceed beyond the stage of 6-phosphogluconate (6-PG), if extracts containing the specific dehydrogenase for this substrate are present together with TPN (Warburg & Christian, 1936, 1937; Lipmann, 1936; Dickens, 1936, 1938a). The end-point of the reaction depends upon the purity of the enzyme extracts; while crude extracts cause extensive oxidation, the somewhat purified dehydrogenase causes consumption of about $\frac{1}{2}$ O_2 per molecule of 6-phosphogluconate, and elimination of about 1 mole of CO_2. Lipmann (1936) suggested that this oxidation, which is not blocked by bromoacetate or by fluoride, might result in the formation of a pentosephosphate, which he thought would be D-arabinose-5-phosphate. Dickens (1938a, 1938b) found, however, that this compound was neither oxidized nor fermented at a rate sufficient for it to be an intermediate in this type of hexose degradation. D-Ribose-5-phosphate, however, he found was entirely suitable in this respect to fill the role of the intermediate. Consequently it was suggested that during the oxidation of 6-phosphogluconate the necessary change of configuration of the sugar molecule occurred at carbon atom 2 of the pentose, giving D-ribose-5-phosphate (Dickens, 1938a, 1938b).

Cohen and Scott took up this problem again more than ten years later, and proved by chromatography and by the use of adaptive strains of micro-organisms that D-ribose was indeed one of the products formed in the same yeast-enzyme system as that originally used by Dickens (Cohen & Scott, 1950; Cohen, Scott & Lanning, 1951; Scott & Cohen, 1951, 1952). These enzyme extracts were then much purified by Horecker & Smyrniotis (1950, 1951), and subjected to detailed study (Horecker, Smyrniotis & Seegmiller, 1951). The first product identified from the action of the specific dehydrogenase on 6-phosphogluconate proved to be laevorotatory and was identified as a phosphorylated derivative of the ketopentose, ribulose (or 2-ketoribose). Presumably this

FIG. 2. THE HEXOSEMONOPHOSPHATE OXIDATIVE ROUTE: DECARBOXYLATIVE TRANSFORMATION OF HEXOSE TO PENTOSE



A : glucose-6-phosphate dehydrogenase
B : 6-phosphogluconic dehydrogenase
(Both A and B require triphosphonicotinamide-adenine dinucleotide (TPN), as coenzyme.)
C : phosphopentose isomerase

ester is ribulose-5-phosphate, which is converted by an enzyme (phosphopentose isomerase) present in the extracts to ribose-5-phosphate, which was also identified. The general course of these reactions is indicated in fig. 2. The primary formation of the ketopentose overcomes the difficulty concerning the isomerization at C_2 of the pentose remarked on earlier by Dickens. The action of the dehydrogenase has been shown by Horecker & Smyrniotis (1952c) to be moderately readily reversible:

$$\text{6-phosphogluconate} + \text{TPN} \xrightleftharpoons{\text{6-PG dehydrogenase}} \text{ribulose-5-phosphate} + CO_2 + \text{TPNH}_2$$

$$\text{ribulose-5-phosphate} \xrightleftharpoons{\text{pentose isomerase}} \text{ribose-5-phosphate}$$

Carbon dioxide labelled with ^{14}C was found to be incorporated in presence of the enzymes into the molecule of phosphogluconate. Thus a reductive synthesis of hexose from pentose and carbon dioxide is rendered biologically feasible by this work of Horecker and his colleagues.

Phosphogluconic dehydrogenase is widely distributed in nature, and generally appears to accompany glucose-6-phosphate dehydrogenase, as already mentioned (Dickens & Glock, 1950, 1951; Horecker & Smyrniotis, 1951, 1952a, 1952b, 1952c). It is a sulphydryl enzyme (Glock & McLean, 1953) which is resistant to halogenacetates.

5. Further Metabolism of Pentosephosphate

If ribose-5-phosphate is incubated anaerobically with a yeast extract, it is almost completely fermented to give 1 mole of ethanol, 1 mole of inorganic phosphate and 1 mole of carbon dioxide; leaving 2 of the 5 carbon atoms unaccounted for (Dickens, 1938b). Evidently cleavage to give a C_2 fragment, accompanied by alcoholic fermentation of the triosephosphate residue, must have occurred (cf. Racker, 1948, 1952). In presence of purified liver extracts (Glock, 1952a, 1952b; Horecker & Smyrniotis, 1952a, 1952b, 1952c) or blood haemolysates (Dische, 1938, 1951; Dische & Pollaczek, 1952; Waldvogel & Schlenk, 1947), the disappearance of pentose is accompanied by the formation of a hexose phosphoric ester. One mode of formation of the latter would obviously be via condensation of 2 moles of triosephosphate by aldolase to give 1 mole of fructose-1:6-diphosphate. It appears, however, that another route to hexose exists, yielding hexosemonophosphate (Robison equilibrium ester) without passing through the intermediate stage of fructosediphosphate (Waldvogel & Schlenk, 1947; Dickens & Glock, 1951). Not only are liver extracts which are deficient in fructosediphosphatase still able to form hexosemonophosphate abundantly, but the yield is in fact considerably in excess of that which could be derived merely from the triose fragment of the pentosephosphate molecules (Glock, 1952a, 1952b). How is the remaining C_2 fragment of the C_5 chain incorporated? This is a puzzle to which only a partial solution can at present be given.

6. Part Played by a C_7 Sugar in the Anaerobic Transformation of Pentose to Hexose

As long ago as 1938, Robison, Macfarlane & Tazelaar described the accumulation of a ketoheptose ester during yeast-juice fermentation. Horecker & Smyrniotis (1952a, 1952b), in studying the orcinol colour reaction during the transformation of pentosephosphate in tissue extracts, found that the light absorption peak at 670 mμ. corresponding to pentose was replaced by one at 580–600 mμ., which corresponded with the reaction for a ketoheptose. These authors used a liver fraction together with added muscle aldolase. Glock (1952b) has also found a similar change in the cysteine-carbazole colour reaction for ketoses when ribose-5-phosphate is incubated with a partially purified enzyme fraction from liver. Horecker & Smyrniotis (1952a) have now been able to identify the ketoheptose as sedoheptulose, which occurs as a phosphoric ester. Various possible routes by which this sugar might arise from pentose are suggested in fig. 3. The earlier work of Horecker and Smyrniotis suggests a cleavage of ribulose-5-phosphate formed as a first product:

$$2(\text{ribose-5-phosphate}) \xrightleftharpoons{\text{isomerase}} 2(\text{ribulose-5-phosphate})$$

$$\text{"splitting enzyme"} \quad 2(C_2 \text{ compound}) + 2(C_3 \text{ compound})$$

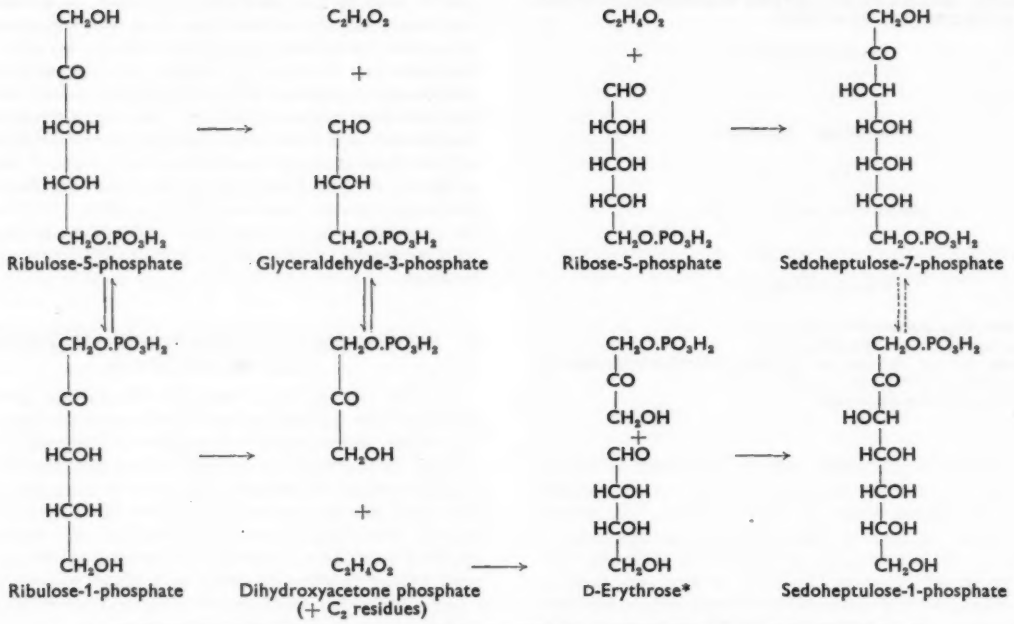
$$\text{"aldolase"} \quad C_7 \text{ compound} + C_3 \text{ compound}$$

According to this formulation one molecule of triose should remain over, and Horecker's latest work appears to confirm that a route of this kind exists (Horecker, 1953). Several routes for condensation of two C_2 units + one C_3 unit (e.g., of the tetrose erythrose + triosephosphate, cf. Horecker & Smyrniotis, 1952a) may be suggested, but evidence showing the exact route of synthesis of sedoheptulose is not yet available.

Another unsolved puzzle is the way in which the seven-carbon sugar is apparently transformed to hexose. For example, Glock (1952b) found that in liver extracts fructose-6-phosphate, which was slowly transformed into Robison-equilibrium hexosemonophosphate, was the first detectable hexose formed from ribose-5-phosphate in her experiments. Possibly C_2 units of a highly reactive type are split off from the molecule of sedoheptulose phosphate and recombine to form the hexose molecule. Some indirect support for this idea is given by the work of Akabori and his colleagues (Akabori & Uehara, 1952; Akabori, Uehara & Muramatsu, 1952), who found that hydroxypyruvic acid (formed in the

UNIVERSITY OF MICHIGAN LIBRARIES

FIG. 3. SEDOHEPTULOSE FORMATION (Horecker, 1953)



* The latest results obtained by Horecker & Smyrniotis (1953), and by Racker, de la Haba & Leder (1953), suggest that ribulose-5-phosphate may be converted to a tetrose (? L-erythulose) by a reaction requiring thiamine pyrophosphate as coenzyme. D-Glyceraldehyde-3-phosphate is also formed, and by condensation with L-erythulose may form sedoheptulosephosphate. (Author's note added in proof.)

enzymic decarboxylation of dihydroxymaleic acid by rabbit tissues) itself undergoes decarboxylation and condensation in tissue extracts yielding first a tetrose and then a hexose. This may be another example of a source of reactive C_2 units capable of synthesizing higher sugars.

However this may be, the exact mechanism of hexose production from pentose, via ketoheptose, remains to be worked out. It is, however, of great interest to recall that sedoheptulose is widely distributed in the vegetable kingdom, and, according to recent work on photosynthesis, sedoheptulose phosphate is an early product of this vital activity of the plant, being apparently quickly converted to hexose-phosphate, in much the same manner as has been suggested above (Benson, Bassham & Calvin, 1951). These results combine to suggest a new and functional importance for the hitherto supposedly obscure ketopentoses and ketoheptoses in both the animal and vegetable kingdoms.

7. Distribution of these Metabolic Pathways in Nature and their Relative Significance

At present we do not know for any cell or tissue just what proportion of the total carbohydrate oxidation proceeds via one or other of the two pathways already described. Probably in brain tissue, for example, iodoacetate- and fluoride-insensitive pathways exist (Barker, Shorr & Malam, 1939), which may well proceed via the hexosemonophosphate oxidative route, when these reagents have blocked the glycolytic pathway. Possibly, even in the absence of these inhibitors, brain respiration may proceed similarly—we do not yet know. Engelhardt & Sakov (1943) favour the view that under aerobic conditions the formation of fructose-diphosphate by phosphorylation of the monophosphate is blocked, and consequently the glycolytic pathway does not function aerobically.

The most direct evidence is that of Cohen (1951) on *Bacterium coli*. As a result of his studies of the metabolism of glucose labelled with ^{14}C, Cohen concludes that about 30–40% of the total respiration of these cells growing in a glucose medium can proceed via the hexosemonophosphate oxidative route. This also agrees with the distribution of the respective enzymes per cell of the organism (Scott & Cohen, 1952).

The fact, described in some detail above, that glucose is continually regenerated by a cyclical process in which C_1 is first eliminated as carbon dioxide, makes it hard to devise really satisfactory methods of assessing the significance of the hexosemonophosphate oxidative route. The investigations to date suggest that its true importance may be much greater than the amount of work at present devoted to this pathway would indicate. This applies not only to its role in energy metabolism but also to its probable significance as a source of essential pentosephosphates in the living organism.

8. Summary

In addition to the enzymes of the well-known pathway of hexose breakdown which occurs by way of the Harden-Young ester, fructose-1:6-diphosphate, many types of animal and vegetable cells contain enzymes able to oxidize hexosemonophosphate directly by TPN-linked dehydrogenases, without passage through the stage of hexosediphosphate. The first product of this oxidation, the lactone of 6-phosphogluconic acid, is also capable of further oxidative decarboxylation, yielding, as first suggested by Dickens (1938a), the biologically important pentosephosphate, ribose-5-phosphate, which is a constituent of many nucleotides, including coenzymes of oxidation and phosphorylation, and of the ribonucleic acids. Recent work from the laboratories of Cohen and Horecker has firmly established this route of biological formation of pentose from hexose, and the latter worker has been able to demonstrate the possibility of reversal of this reaction, which may have special importance in photosynthesis. Horecker and his colleagues have also shown that the primary product formed is the ketopentose ester, ribulosephosphate, which can be enzymically transformed into ribose-5-phosphate.

Further recently described enzyme systems are involved in the further metabolism of pentosephosphate. Ribose-5-phosphate was originally shown by Dickens (1938b) to be fermented by yeast enzymes in presence of DPN, giving alcohol, carbon dioxide, inorganic phosphate and an unidentified product of the two-carbon residue. This showed that triosephosphate and pyruvate can be produced enzymically from ribose-5-phosphate. Recent work from British and USA laboratories confirms this, and also shows that again an additional route occurs in this transformation, by means of which hexosemonophosphate can be resynthesized from pentosephosphate without passing through the stage of fructose-1:6-diphosphate. This route yields more hexose than could be obtained by the simple union, by means of aldolase, of two triose fragments derived from two molecules of pentose. It appears to involve the intermediate stage of a seven-carbon ketose ester, sedoheptulosemonophosphate. Although the detailed mechanism of this reaction is not yet worked out, it is of great interest that sedoheptulose, which has hitherto been described as a fairly widespread plant product, and has recently been found to participate in photosynthesis, should now be found as a product of animal metabolism also.

Incomplete as knowledge of these systems still is, there can be little doubt that they have important functional significance, not only in the plant and micro-organism but also in animal physiology. They may provide alternatives to the fermentative or glycolytic pathway which can at least partially replace it in aerobic carbohydrate metabolism. They are also oxidative sources of essential pentosephosphate, additional to the synthesis of this compound by condensation of C_3 and C_2 fragments. An attempt has been made to estimate the relative importance of the above two routes in this respect, but far more data are necessary on this subject. It must always be remembered that still other pathways of carbohydrate oxidation, for example, via gluconic acid (cf. Stetten & Stetten, 1950) or via glucuronic acid (cf. Douglas & King, 1952), also exist and may have at any rate a limited significance as alternative pathways of carbohydrate metabolism.

References

Akabori, S. & Uehara, K. (1952) *II^e Congrès International de Biochimie, Paris, 21–27 Juillet 1952: Résumés des communications*, p. 30. Société d'Edition d'Enseignement Supérieur, Paris

Akabori, S., Uehara, K. & Muramatsu, I. (1952) *Proc. imp. Acad. Japan*, **28**, 39

Barker, S. B., Shorr, E. & Malam, M. (1939) *J. biol. Chem.* **129**, 33

Benson, A. A., Bassham, J. A. & Calvin, M. (1951) *J. Amer. chem. Soc.* **73**, 2970

Cohen, S. S. (1951) *Nature, Lond.* **168**, 746

Cohen, S. S. & Scott, D. B. McNair (1950) *Science*, **111**, 543

Cohen, S. S., Scott, D. B. McNair & Lanning, M. (1951) *Fed. Proc.* **10**, 173

Cori, O. & Lipmann, F. (1952) *J. biol. Chem.* **194**, 417

Dickens, F. (1936) *Nature, Lond.* **138**, 1057

Dickens, F. (1938a) *Biochem. J.* **32**, 1626

Dickens, F. (1938b) *Biochem. J.* **32**, 1645

Dickens, F. & Glock, G. (1950) *Nature, Lond.* **166**, 33

Dickens, F. & Glock, G. (1951) *Biochem. J.* **50**, 81

Dische, Z. (1938) *Naturwissenschaften*, **26**, 252

Dische, Z. (1951) In: McElroy, W. D. & Glass, H. B., ed. *Phosphorus metabolism: a symposium on the role of phosphorus in the metabolism of plants and animals*, vol. 1, p. 171. Johns Hopkins Press, Baltimore

Dische, Z. & Pollaczek, E. (1952) *II^e Congrès International de Biochimie, Paris, 21–27 Juillet 1952: Résumés des communications*, p. 289. Société d'Edition d'Enseignement Supérieur, Paris

Douglas, J. F. & King, C. G. (1952) *J. biol. Chem.* **198**, 187

Engelhardt, V. A. & Sakov, N. E. (1943) *Biochimica*, **8**, 9

Gibbs, M. (1952) *Nature, Lond.* **170**, 164

Glock, G. E. (1952a) *Nature, Lond.* **170**, 162

Glock, G. E. (1952b) *Biochem. J.* **52**, 575

Glock, G. E. & McLean, P. (1952) *Nature, Lond.* **170**, 113

Glock, G. E. & McLean, P. (1953) *Biochem. J.* **55** (In press)

Harden, A. & Young, W. J. (1909) *Proc. roy. Soc.* B, **81**, 336

Horecker, B. L. (1951) In: McElroy, W. D. & Glass, H. B., ed. *Phosphorus metabolism: a symposium on the role of phosphorus in the metabolism of plants and animals*, vol. 1, p. 117. Johns Hopkins Press, Baltimore

Horecker, B. L. (1953) *J. cell. comp. Physiol.* Suppl. (In press)

Horecker, B. L. & Smyrniotis, P. Z. (1950) *Arch. Biochem.* **29**, 232

Horecker, B. L. & Smyrniotis, P. Z. (1951) *J. biol. Chem.* **193**, 371

Horecker, B. L. & Smyrniotis, P. Z. (1952a) *J. Amer. chem. Soc.* **74**, 2123

Horecker, B. L. & Smyrniotis, P. Z. (1952b) *Fed. Proc.* **11**, 232

Horecker, B. L. & Smyrniotis, P. Z. (1952c) *J. biol. Chem.* **196**, 135

Horecker, B. L. & Smyrniotis, P. Z. (1953) *J. Amer. chem. Soc.* **75**, 1009

Horecker, B. L., Smyrniotis, P. Z. & Seegmiller, J. E. (1951) *J. biol. Chem.* **193**, 383

Krebs, H. A. (1943) *Advanc. Enzymol.* **3**, 191

Ley, J. de & Cornut, S. (1951) *Nature, Lond.* **168**, 515

Lipmann, F. (1936) *Nature, Lond.* **138**, 588

Peters, R. A. (1936) *Biochem. J.* **20**, 2206

Racker, E. (1948) *Fed. Proc.* **7**, 180

Racker, E. (1952) *J. biol. Chem.* **196**, 347

Racker, E., Haba, G. de la & Leder, I. G. (1953) *J. Amer. chem. Soc.* **75**, 1010

Robison, R. (1922) *Biochem. J.* **16**, 809

Robison, R., Macfarlane, M. G. & Tazelaar, A. (1938) *Nature, Lond.* **142**, 114

Schlenk, F. (1951) In: Sumner, J. B. & Myrbäck, K. D. R., ed. *The enzymes: chemistry and mechanism of action*, vol. 2, pt 1, p. 260. Academic Press, New York

Scott, D. B. McNair & Cohen, S. S. (1951) *J. biol. Chem.* **188**, 509

Scott, D. B. McNair & Cohen, S. S. (1952) *Fed. Proc.* **11**, 284

Seegmiller, J. E. & Horecker, B. L. (1951) *J. biol. Chem.* **194**, 261

Stetten, M. R. & Stetten, DeW., jr (1950) *J. biol. Chem.* **187**, 241

Waldvogel, M. J. & Schlenk, F. (1947) *Arch. Biochem.* **14**, 484

Warburg, O. & Christian, W. (1936) *Biochem. Z.* **287**, 440

Warburg, O. & Christian, W. (1937) *Biochem. Z.* **292**, 287

Warburg, O., Christian, W. & Griese, A. (1935) *Biochem. Z.* **282**, 157

UNIVERSITY OF MICHIGAN LIBRARIES

PHOSPHATE-TRANSFERRING ENZYMES

M. DIXON Sc.D. F.R.S.

Reader in Enzyme Biochemistry, University of Cambridge

E. C. WEBB Ph.D.

University Lecturer in Biochemistry
University of Cambridge

1 Phosphokinases
2 Phosphomutases
3 " Phosphotransferases "
References

The existence of enzymes which hydrolyse organic phosphates to inorganic phosphate has been known for many years, but the discovery of enzymes which bring about the direct transfer of the phosphate radical from one molecule to another, without the intermediate formation of free phosphate, is much more recent. It now appears that the reactions catalysed by the enzymes attacking the phosphate group in organic phosphate compounds fall into three categories: (i) hydrolysis to free phosphate (phosphatases)[1]; (ii) phosphorolysis, i.e., the splitting of a molecule into two parts with the insertion of a phosphate group (phosphorylases); (iii) transfer of the phosphate radical from one molecule to another.

This review is concerned only with the last type, which is brought about by three distinct groups of enzymes: *phosphokinases*, which catalyse the transfer of phosphate from adenosinetriphosphate (ATP) to organic molecules; *phosphomutases*, which appear to transfer a phosphate group from one position to another in the same molecule, but have in fact recently been shown to act by transferring the phosphate from one molecule to another; and "*phosphotransferases*", which do not involve ATP, but catalyse phosphate transfer between a variety of different molecules.

The work of the last decade has shown the great biological importance of phosphate compounds and the enzymes acting on them.

1. Phosphokinases

The name "phosphokinase" was introduced by Dixon and Needham in 1940 (see Dixon & Needham, 1946) to denote enzymes which catalyse reactions of the type

$$ATP + X = ADP + XP$$

by direct transfer of the terminal phosphate group of ATP to the organic molecule X, yielding the phosphorylated compound (XP) and adenosinediphosphate (ADP). The first of these enzymes to be discovered was hexokinase (X = glucose) (Meyerhof, 1927), but the first elucidation of the nature of the phosphokinase type of reaction was due to Lohmann (1934) for creatine phosphokinase, and to Parnas, Ostern & Mann (1934) for pyruvate phosphokinase. The majority of phosphokinases are of comparatively recent discovery; a list is given in Table I, which also includes abbreviations adopted.

The phosphokinases play an important part in biosynthesis. It is now generally believed that the available energy from biological oxidations is first used for the synthesis of pyrophosphate bonds in ATP. The function of the phosphokinases is to make this energy available for organic syntheses by phosphorylating organic compounds, which can then be linked with other molecules by splitting out the phosphate radical.

The phosphokinases all seem to have the common property that they require the presence of Mg^{++} for their activity. All those that have been tested have been found to be "—SH enzymes", i.e., to depend for their activity on sulphydryl groups within their molecules, and it is probable that this also is a general property.

Hexokinases. These enzymes catalyse the first stage in the normal metabolism of the free sugars. The mechanism of the reaction (i.e., ATP + glucose = ADP + glucose-6-phosphate) was first elucidated by von Euler & Adler (1935) in the case of yeast hexokinase. This enzyme was crystallized by Berger *et al.* (1946), by Kunitz & MacDonald (1946) and by Bailey & Webb (1948); all these purifications made use of the marked stabilizing action of glucose on the enzyme which was discovered by R. van Heyningen (1941). Interest in the enzyme had been aroused by war-time studies which showed a parallelism between vesicant properties and power to inhibit hexokinase (Dixon & Needham, 1946).

The enzymes of animal tissues have not been crystallized. There is some difficulty in obtaining active extracts from such tissues as muscle and liver, and, while fresh extracts show hexokinase activity, aged extracts fail to catalyse glucolysis without the addition of hexokinase.

Yeast hexokinase will catalyse the phosphorylation of glucose, fructose and mannose (van Heyningen, 1941; Kunitz & MacDonald, 1946) but not of nine other sugars tested. ADP is inactive in place of ATP. Gottschalk (1941, 1947), in experiments at 0° C., showed that α- and β-glucose and α- and β-mannose are all readily attacked; fructose is attacked only in the furanose form. D-Glucosamine is phosphorylated almost as readily as glucose (Brown, 1951; Grant & Long, 1951).

The enzyme is irreversibly inhibited by —SH reagents (e.g., chloroacetophenone), against which it can be protected by arsenicals which combine reversibly with —SH groups (van Heyningen, see Dixon, 1948).

Recrystallized yeast hexokinase attacks glucose, fructose and mannose, and studies on mutual competition have shown convincingly that one and the same enzyme is responsible for all three reactions (Slein, Cori & Cori, 1950). The same reactions and competition effects were obtained with hexokinase from brain, which also phosphorylates D-glucosamine (Harpur & Quastel, 1949), and it appears that this enzyme closely resembles the yeast enzyme. Crude tumour hexokinase also rapidly attacks all three sugars (Boyland, Goss & Williams-Ashman, 1951). Inhibition of brain hexokinase by the products of the reaction was studied by Weil-Malherbe & Bone (1951a), who suggested a mechanism of action.

On the other hand, in the case of muscle and liver, evidence has been produced that glucose and fructose are phosphoryl-

[1] See King, p. 160 of this Bulletin.

TABLE I. PHOSPHOKINASES

Type of phosphoric compound (XP) formed	Phosphate acceptor (X)	Compound formed (XP)	Special name of enzyme (if any)
Low-energy esters	Glucose, Fructose, Mannose, Glucosamine	Hexose-6-phosphate	Hexokinase (yeast, brain)
	Glucose	Glucose-6-phosphate (?)	Glucokinase (muscle, liver)
	Fructose, Other ketoses	Fructose-1-phosphate, etc.	Ketohexokinase (muscle, liver)
	Ribose	Ribosephosphate	Ribokinase
	Fructose-6-phosphate	Fructose-1:6-diphosphate	6-Phosphofructokinase
	Fructose-1-phosphate	Fructose-1:6-diphosphate	1-Phosphofructokinase (muscle)
	Glucose-1-phosphate	Glucose-1:6-diphosphate	Phosphoglucokinase
	Gluconate	6-Phosphogluconate	Gluconokinase
	Adenosine	Adenylate (AMP)	—
	Riboflavin	Flavin nucleotide	Flavokinase
	Nicotinamide nucleoside	Nicotinamide nucleotide	
	Coenzyme I	Coenzyme II	—
	Pyridoxal	Codecarboxylase	—
	Protein	Phosphoprotein	—
Glycosidic phosphates	Galactose	Galactose-1-phosphate	Galactokinase
High-energy phosphates	Adenylate (AMP)	Adenosinediphosphate (ADP)	Myokinase
	3-Phosphoglycerate	1:3-Diphosphoglycerate	—
	Acetate	Acetylphosphate	Acetokinase
	Pyruvate (Pyr)	Phosphoenolpyruvate (PyrP)	—
	Creatine (Cr)	Phosphocreatine (CrP)	—
	Arginine	Phosphoarginine	—
	Glutamate (?)	γ-Glutamylphosphate (?)	—
	Luciferin (?)	?	—
Pyrophosphates	Coenzyme A	Coenzyme A pyrophosphate	Coenzyme A pyrophosphokinase
	Thiamine	Cocarboxylase	Thiamine pyrophosphokinase

ated by different enzymes (Cori & Slein, 1947; Cori *et al.* 1951) which were obtained separately in different protein fractions.

The specificity of glucokinase has not yet been examined and it has not been definitely shown that the product is glucose-6-phosphate. The rates of phosphorylation of glucose in different animal tissues have been determined by Long (1951), who found that brain was by far the most active tissue.

The fructokinase of liver and muscle has been partially purified (Slein *et al.* 1950; Leuthardt & Testa, 1950a) and the product shown to be fructose-1-phosphate (Cori *et al.* 1951; Leuthardt & Testa, 1951). It phosphorylates not only fructose, but also the other ketoses D-sorbose and D-tagatose; it does not act on glucose, galactose or mannose (Leuthardt & Testa, 1950b). The enzyme has therefore been called "ketohexokinase". According to Slein *et al.* (1950) the affinity of the muscle enzyme for fructose is immeasurably low; but according to Hers (1952) that of the liver enzyme is immeasurably high.

Studies on the corresponding bacterial systems have shown that glucose and fructose are phosphorylated by different enzymes. Klein & Doudoroff (1950) have studied the specific glucokinase of *Pseudomonas putrefaciens*; Cardini (1951) separated the glucokinase of *Bacterium coli* from fructokinase, but found that it still acted on mannose. *Staphylococcus aureus* glucokinase, on the other hand, has no activity towards either fructose or mannose.

The Cori school have shown that hormones have a marked action on hexokinase and they believe that their action on carbohydrate metabolism is produced in this way. The enzyme is strongly inhibited in vitro by extracts of anterior pituitary (Colowick, Cori & Slein, 1947); and this has been confirmed by other workers, e.g., Reid, Smith & Young (1948). According to Colowick and his colleagues the

UNIVERSITY OF MICHIGAN LIBRARIES

activity is restored by the addition of insulin in a number of cases, but this has not been confirmed by Stadie & Zapp (1947).

Weil-Malherbe & Bone (1951b) have obtained a specific activator of hexokinase, non-dialysable and relatively heat-stable, from erythrocytes and from rabbit-muscle extract. The activator from muscle has been partially purified and is particularly active with brain hexokinase.

Galactokinase differs from the other hexokinases in forming a glycosidic phosphate (Trucco *et al.* 1948). It is specific for galactose and will not act on glucose, fructose or mannose. It occurs both in animal tissues and in micro-organisms.

A ribokinase has been partially purified from baker's yeast by Sable (1950). The product of the reaction was not identified. The preparation did not act on D-xylose, D- or L-arabinose or D-2-deoxyribose.

Phosphohexokinases. These enzymes bring about the further phosphorylation of phosphohexoses, giving diphosphohexoses. The 6-phosphofructokinase is present in yeast and in animal tissues, but only in low concentration in liver. The enzyme from rabbit muscle has been partially purified by Taylor (1947) and by Racker (1947). It does not act on glucose or fructose. The enzyme from brain has been partially purified bv Muntz (1952) and found not to act on glucose, fructose, glucose-1-phosphate or glucose-6-phosphate.

A 6-phosphohexokinase has also been demonstrated in plants and purified sixfold from pea-meal by Axelrod *et al.* (1952). Their preparation does not act on glucose, fructose or glucose-1-phosphate.

Muscle contains another phosphohexokinase which phosphorylates fructose-1-phosphate in the 6-position (Slein *et al.* 1950). Muscle and yeast contain yet another enzyme of this type which phosphorylates glucose-1-phosphate in the 6-position (Paladini *et al.* 1949). This enzyme is of importance because it produces glucose-1:6-diphosphate, the "coenzyme" of phosphoglucomutase (see section 2).

Gluconokinase. This enzyme has been partially purified from *Bact. coli* (Cohen, 1950) and from yeast (Sable & Guarino, 1951) and has been separated from adenosinetriphosphatase, hexokinase and the other glycolytic enzymes.

Nucleoside Phosphokinases. Adenosine phosphokinase catalyses the phosphorylation of adenosine to adenosine-5′-phosphate (adenylic acid). Caputto (1951) showed the enzyme to be present in yeast, kidney and liver extracts. Kornberg & Pricer (1951) purified the yeast enzyme and showed it to be specific for adenosine and 2-aminoadenosine among the 17 nucleotides tested.

A nicotinamide nucleoside phosphokinase, forming nicotinamide nucleotide from nicotinamide riboside in the presence of ATP, was shown to be present in a protein fraction from pig liver by Rowen & Kornberg (1951). It thus brings about one stage of the enzymic synthesis of coenzymes I and II.

The corresponding step in the formation of the flavin nucleotides is catalysed by flavokinase (Kearney & Englard, 1951a, 1951b; Kearney, 1952). The enzyme was purified 1,200 times from brewer's yeast and its specificity studied. It does not act on adenosine, ribose, *iso*riboflavin, galactoflavin, dulcitylflavin or sorbitylflavin, but phosphorylates riboflavin, arabitylflavin and 6:7-dichlororiboflavin, all in the 5′-position. Unlike other phosphokinases, it can utilize ADP as well as ATP, though more slowly, despite the absence of myokinase. It was shown by Englard (1952) to be present also in extracts of rat intestine.

Other Phosphokinases of this Group. Coenzyme I is converted into coenzyme II (2′-phospho-coenzyme I) by a phosphokinase which has been partially purified from brewer's yeast by Kornberg (1950). The phosphorylation occurs in the ribose adjacent to the adenine.

The formation of pyridoxal phosphate (codecarboxylase) from pyridoxine and ATP in liver homogenates was shown by Binkley & Christensen (1951). A thorough investigation of the enzyme in brewer's yeast which phosphorylates pyridoxal was made by Hurwitz (1952). The enzyme, purified 1,000 times, acted also on pyridoxine, pyridoxamine and various substituted pyridoxines. The products from these substances act as competitive inhibitors of codecarboxylase.

The existence of a phosphokinase bringing about the phosphorylation of protein has been noted in the yolk of developing frog eggs by Barth & Jaeger (1950).

Phosphokinases Forming High-Energy Phosphates. The enzymes mentioned above form low-energy phosphate compounds from the high-energy ATP. The free energies (and therefore the equilibrium constants) of the reactions are consequently so large that they are effectively irreversible. We come now to the phosphokinases which form high-energy phosphate compounds, i.e., compounds in which the phosphate bond energy is of the same order as that of ATP. The free energies of the reactions are therefore comparatively small and they are consequently readily reversible.

Myokinase (adenylate phosphokinase) catalyses the phosphorylation of adenylic acid (adenosinemonophosphate, AMP) to ADP, and thus produces two molecules of ADP, one from AMP and one from ATP:

$$\text{AMP} + \text{ATP} \rightleftharpoons 2\text{ADP}.$$

The phosphokinases in general utilize only the terminal phosphate group of ATP, but in the presence of myokinase the second group also becomes available through the formation of ATP from the ADP produced.

Myokinase was discovered in this way in connexion with hexokinase (Colowick & Kalckar, 1943; Kálckar, 1943). The enzyme from skeletal muscle has the remarkable property of being unaffected by heating to 100° C. in 0.1 N-hydrochloric acid for several minutes. Yeast and liver myokinases, however, are destroyed under these conditions (Kotel'nikova, 1948). The equilibrium constant of the myokinase reaction has recently been accurately determined by Eggleston & Hems (1952). At 25° C. and pH 7.4 they find that at equilibrium $\frac{[\text{ATP}]\,[\text{AMP}]}{[\text{ADP}]^2} = 0.444$; in other words the energies of the first and second pyrophosphate links in ATP differ by only a few hundred calories.

Creatine phosphokinase is responsible for the Lohmann reaction (Lohmann, 1934):

$$\text{phosphocreatine} + \text{ADP} \rightleftharpoons \text{creatine} + \text{ATP}.$$

It is particularly active in muscle, in which it enables the creatine phosphate present to act as a reservoir of available energy. Preparations of the enzyme which appear to be at least 85% pure were obtained by Askonas (1951a), who studied the equilibrium constants of the reaction. The equilibrium depends upon the pH (Lehmann, 1936; Dixon, 1949), but it also depends markedly upon the nature and concentration of the activating metal ion, e.g., Mg^{++} (Askonas, 1951a). This important effect, which may be of general

occurrence in phosphate-transfer reactions, is due to the fact that metal ions form complexes with the pyrophosphate groups in ATP and ADP. As ATP contains two pyrophosphate links and ADP only one, the equilibrium would be expected to depend on the Mg concentration. It is not possible to reach equilibrium in the complete absence of Mg, as the enzyme is then inactive, but on progressively reducing the Mg concentration the equilibrium constant $\left(\frac{[ATP]\ [Cr]}{[ADP]\ [CrP]}\right)$ falls and tends towards the true value. By extrapolation to zero concentration Askonas found a value of 1.5 at pH 8.5.

Askonas (1951b) found that this enzyme is strongly inhibited in vitro by low concentrations of thyroxine, while the physiologically inactive diiodotyrosine had no effect.

Evidence of the existence of a corresponding enzyme in invertebrate muscle producing arginine phosphate was obtained by various workers (Lohmann, 1935; Lehmann, 1936; Baldwin & Needham, 1937).

The phosphoglycerate phosphokinase of yeast was crystallized by Bücher (1947) and found to have an extremely high activity. It is also present in animal tissues. It is responsible for feeding the energy of glycolysis and fermentation into the ATP system. It is completely specific. Studies on the mutual inhibition of the reactants led Bücher to suggest that ADP and 3-phosphoglycerate are attached to two different points at the active centre of the enzyme, and the phosphate group to be transferred is attached to the enzyme at a point between them.

Acetokinase (Stern & Ochoa, 1951) occurs in extracts of *Bact. coli*, but not in animal tissue or yeast extracts. The transacetylase, which is also present, can be destroyed by heating to 60° C., leaving the acetokinase.

The pyruvate phosphokinase of human muscle was crystallized by Kubowitz & Ott (1943). Meyerhof & Oesper (1949) found the equilibrium constant of the reaction $\left(\frac{[ATP]\ [Pyr]}{[ADP]\ [PyrP]}\right)$ to be about 2,000 at 30° C. (The pH was not stated, although the constant is known to vary with pH.)

A phosphokinase may be involved in glutamine formation from glutamate and ATP (Speck, 1949; Elliott, 1948, 1951), but the probable intermediate, γ-glutamylphosphate, has not yet been demonstrated.

Another phosphokinase may be involved in bioluminescence, since in the presence of luciferase the addition of ATP to firefly luciferin causes luminescence (McElroy, 1951; see also Harvey & Haneda, 1952).

Pyrophosphokinases. There seem to be at least two enzymes which transfer not one but two phosphate groups from ATP as a pyrophosphate radical. Weil-Malherbe (1939) showed that a soluble enzyme from yeast produced cocarboxylase (thiamine pyrophosphate) from thiamine and ATP. He suggested a direct transfer of pyrophosphate, rather than two successive transfers of phosphate, since thiamine monophosphate is utilized more slowly than thiamine, and he considered that the monophosphate acted only because it was first hydrolysed to thiamine. The phosphokinase was partially purified by van Thoai & Chevillard (1949) and further by Steyn-Parvé (1952), who confirmed Weil-Malherbe's suggestion. The enzyme is also present in animal tissues.

Lipmann *et al.* (1952) have shown that the "active acetyl intermediate" of metabolism (acetyl-coenzyme A) is formed by the elimination of pyrophosphate from coenzyme A pyrophosphate in the presence of acetate. The coenzyme A pyrophosphate is formed from coenzyme A and ATP by a pyrophosphokinase, which they obtained from yeast extract and also found in pigeon-liver extract.

2. Phosphomutases

These phosphate-transferring enzymes do not involve ATP, but appear to move phosphate groups from one position to another in the same molecule, as for instance in the transformation of glucose-1-phosphate into glucose-6-phosphate. Actually, however, the transfer is from position 1 in one molecule to position 6 in another molecule, thus (Caputto *et al.* 1948):

glucose-1-phosphate + glucose-1:6-diphosphate =
glucose-1:6-diphosphate + glucose-6-phosphate.

The reaction must be started by the addition of a trace of the diphosphate, but as one molecule of this appears for every molecule disappearing, it remains constant in amount and acts catalytically as a "coenzyme" of the reaction. All the phosphomutases appear to act in a similar way; in each case the reaction is a transfer in which "coenzyme" is converted into product and substrate into "coenzyme". This mechanism has been confirmed by Sutherland *et al.* (1949) using isotopes.

Phosphoglucomutase. This enzyme catalyses the above reaction, by transferring a phosphate group from the 1-position of the diphosphate to the 6-position of the glucose-1-phosphate. The enzyme from muscle has been crystallized by Najjar (1948). It does not act on fructose-1-phosphate, but converts mannose-1-phosphate into mannose-6-phosphate (Leloir, 1951) and also ribose-1-phosphate into ribose-5-phosphate (Klenow & Larsen, 1952). The latter reaction occurs at only one-hundredth of the rate of the glucose reaction, ribose-1:5-diphosphate being apparently the "coenzyme". The ribose reaction, however, can be started by the addition of glucose-1:6-diphosphate, which probably gives rise to the ribosediphosphate by the reaction

glucose-1:6-diphosphate + ribose-1-phosphate =
glucose-6-phosphate + ribose-1:5-diphosphate.

Stickland (1949) made the remarkable observation that the phosphoglucomutase is active only in the presence of both Mg^{++} and a second metal ion, e.g., Al^{+++} or Cr^{+++}.

Phosphodeoxyribomutase. Some evidence for the existence of this enzyme in rat liver and calf thymus has been obtained by Manson & Lampen (1951). It is believed to catalyse the conversion of deoxyribose-1-phosphate to the 5-phosphate.

Phosphoglyceromutase. This enzyme catalyses the reversible transformation of 3-phosphoglycerate to 2-phosphoglycerate (Meyerhof & Kiessling, 1935). Sutherland, Posternak & Cori (1950) showed that the reaction takes place through the "coenzyme" 2:3-diphosphoglycerate, the phosphate transfer being from the 3-position of this to the 2-position of the 3-phosphoglycerate. This reaction, like that catalysed by phosphoglucomutase, is on the main line of carbohydrate metabolism.

Diphosphoglyceromutase. This enzyme, which has been obtained from rabbit erythrocytes by Rapoport & Luebering (1952), differs slightly from the other phosphomutases, in

that the " coenzyme " is a monophosphate. It catalyses the reaction

1:3-diphosphoglycerate + 3-phosphoglycerate =
3-phosphoglycerate + 2:3-diphosphoglycerate

by transfer from the 1-position of the 1 : 3-diphosphate to the 2-position of the 3-phosphate. The 1 : 3-diphosphate is formed as an intermediate in carbohydrate metabolism and is converted by this enzyme into the " coenzyme " of phosphoglyceromutase.

Glucose-1-phosphate Mutase. This enzyme, though not a phosphomutase, catalyses a phosphate transfer resulting in a dismutation of two molecules of the 1-phosphate, thus:

glucose-1-phosphate + glucose-1-phosphate =
glucose-1:6-diphosphate + glucose

(Leloir *et al.* 1949). It thus produces the " coenzyme " of phosphoglucomutase. The enzyme occurs in bacteria, but not in animal tissues or yeast. It has been partially purified from *Bact. coli.*

3. " Phosphotransferases "

These are enzymes which, like the phosphomutases, do not depend upon ATP, but which differ from them in transferring phosphate groups between molecules of different kinds.

Axelrod (1948a) observed a transfer of phosphate from aryl phosphates to aliphatic alcohols in preparations of acid phosphatase. He was unable to separate the enzyme responsible from the phosphatase. Appleyard (1948) had independently observed a transfer from phenolphthalein phosphate by acid phosphatase preparations, but ascribed it to a separate enzyme. Axelrod (1948b), using ^{32}P, showed that free phosphate was not an intermediate.

Meyerhof & Green (1950) obtained transfers with a purified intestinal alkaline phosphatase from various high-energy phosphates. In their experiments, unless inorganic phosphate were added, there was no accumulation of the product of transfer, which they explained by its hydrolysis by the phosphatase. From inorganic phosphate and the acceptor alcohol alone there was some formation of the phosphorylated product (as would be expected from a simple reversal of the hydrolytic reaction in accordance with the law of mass action), but on addition of the high-energy phosphate (creatine phosphate) the rate of formation was greatly increased. They interpreted this increase as meaning that the product was formed in two ways, both catalysed by the phosphatase itself. One way was by simple reversal of the hydrolysis; the other way was by a direct transfer of phosphate from the high-energy phosphate. Green & Meyerhof (1952) extended these observations to an acid phosphatase, with which they obtained a larger amount of the phosphorylated product than could be formed by the reversal of hydrolysis. As the enzyme preparations used in this work were not completely pure, the possibility that the transfer reactions were due to an enzyme distinct from the phosphatase was not excluded.

Morton (1952), in an extensive study, using very highly purified preparations of milk and intestinal alkaline phosphatases, produced evidence that the phosphatase itself was responsible for the transfer reactions. Unlike Meyerhof, Morton obtained accumulation of considerable amounts of the product without added inorganic phosphate. The property of transfer appears to be general with the unspecific phosphatases; but Morton found no transfer with specific phosphatases, e.g., adenosine-5-phosphatase or hexosediphosphatase. He also found that transfer could occur not only from high-energy to low-energy positions, but also between compounds of similar energy, e.g., from glycerol to glucose. ATP and ADP are inactive as phosphate donors and are not hydrolysed by the purified phosphatase; AMP, which is hydrolysed, acts also as a donor, but only with the unspecific phosphatase.

In addition to creatine phosphate, aryl phosphates, AMP, and glycerolphosphate, Morton found that phosphopyruvate and hexosediphosphate could act as donors; the best acceptors were glucose, glycerol and propylene glycol; fructose acted to some extent but sucrose was not phosphorylated. Serine, threonine and ethanolamine were phosphorylated, but the rate was not determined. The fact that sucrose did not act shows that not all hydroxyl groups are effective as acceptors. Morton found that α-glycerolphosphate is formed by the phosphate transfer from β-glycerolphosphate to glycerol; alkaline phosphatase can therefore act as a " phosphoglycerolmutase ".

The proportion of the phosphate removed from the donor which was transferred to the acceptor depended on the acceptor concentration; as this is increased the total rate of breakdown of the donor remains constant, but more of the phosphate is transferred to the acceptor and less released as inorganic phosphate. Thus the proportion transferred rises to a limiting value and thereafter remains constant in spite of the fact that in very high acceptor concentrations the total rate of breakdown is diminished. The proportion transferred appeared to be independent of the phosphate bond energy of the reactants. Morton's interpretation is that there is a competition at the enzyme between the acceptor alcohol and water for the phosphate from the donor: if the latter reacts with water, hydrolysis to inorganic phosphate results; if with the acceptor, phosphate transfer takes place.

As a result of this work it is now clear that it is not necessary to assume the existence of a separate class of phosphotransferase enzymes, since the reactions can be accounted for by the ordinary phosphatases.

REFERENCES

Appleyard, J. (1948) *Biochem. J.* **42,** 596
Askonas, B. A. (1951a) *The separation of enzymes by means of organic solvents at low temperature. Application to aqueous rabbit muscle extract, with a study of creatinephosphokinase.* (Thesis for Ph.D. degree.) University of Cambridge
Askonas, B. A. (1951b) *Nature, Lond.* **167,** 933
Axelrod, B. (1948a) *J. biol. Chem.* **172,** 1
Axelrod, B. (1948b) *J. biol. Chem.* **176,** 295
Axelrod, B., Saltman, P., Bandurski, R. S. & Baker, R. S. (1952) *J. biol. Chem.* **197,** 89
Bailey, K. & Webb, E. C. (1948) *Biochem. J.* **42,** 60
Baldwin, E. & Needham, D. M. (1937) *Proc. roy. Soc.* B, **122,** 197
Barth, L. G. & Jaeger, L. (1950) *J. cell. comp. Physiol.* **35,** 413
Berger, L., Slein, M. W., Colowick, S. P. & Cori, C. F. (1946) *J. gen. Physiol.* **29,** 379
Binkley, F. & Christensen, G. M. (1951) *J. Amer. chem. Soc.* **73,** 3535
Boyland, E., Goss, G. C. L. & Williams-Ashman, H. G. (1951) *Biochem. J.* **49,** 321
Brown, D. H. (1951) *Biochim. biophys. Acta,* **7,** 487

Bücher, T. (1947) *Biochim. biophys. Acta*, **1**, 292
Caputto, R. (1951) *J. biol. Chem.* **189**, 801
Caputto, R., Leloir, L. F., Trucco, R. E., Cardini, C. E. & Paladini, A. C. (1948) *Arch. Biochem.* **18**, 201; **19**, 339
Cardini, C. E. (1951) *Enzymologia*, **14**, 362
Cohen, S. S. (1950) *J. biol. Chem.* **189**, 617
Colowick, S. P., Cori, G. T. & Slein, M. W. (1947) *J. biol. Chem.* **168**, 583
Colowick, S. P. & Kalckar, H. M. (1943) *J. biol. Chem.* **148**, 117
Cori, G. T., Ochoa, S., Slein, M. W. & Cori, C. F. (1951) *Biochim. biophys. Acta*, **7**, 304
Cori, G. T. & Slein, M. W. (1947) *Fed. Proc.* **6**, 246
Dixon, M. (1948) *Symp. biochem. Soc.* No. 2, p. 38
Dixon, M. (1949) *Multi-enzyme systems*, p. 98. University Press, Cambridge
Dixon, M. & Needham, D. M. (1946) *Nature, Lond.* **158**, 432
Eggleston, L. V. & Hems, R. (1952) *Biochem. J.* **52**, 156
Elliott, W. H. (1948) *Nature, Lond.* **161**, 128
Elliott, W. H. (1951) *Biochem. J.* **49**, 106
Englard, S. (1952) *Fed. Proc.* **11**, 208
Euler, H. von & Adler, E. (1935) *Hoppe-Seyl. Z.* **235**, 122
Gottschalk, A. (1941) *Aust. J. exp. Biol. med. Sci.* **22**, 291
Gottschalk, A. (1947) *Biochem. J.* **41**, 276
Grant, P. T. & Long, C. (1951) *Biochem. J.* **50**, *Proc.* xx
Green, H. & Meyerhof, O. (1952) *J. biol. Chem.* **197**, 347
Harpur, R. P. & Quastel, J. H. (1949) *Nature, Lond.* **164**, 693
Harvey, E. N. & Haneda, Y. (1952) *Arch. Biochem. Biophys.* **35**, 470
Hers, H. G. (1952) *Biochim. biophys. Acta*, **8**, 466
Heyningen, R. van (1941) In: Dixon, M., Report No. 10 to Ministry of Supply*
Hurwitz, J. (1952) *Biochim. biophys. Acta*, **9**, 496
Kalckar, H. M. (1943) *J. biol. Chem.* **148**, 127
Kearney, E. B. (1952) *J. biol. Chem.* **194**, 747
Kearney, E. B. & Englard, S. (1951a) *Arch. Biochem. Biophys.* **32**, 222
Kearney, E. B. & Englard, S. (1951b) *J. biol. Chem.* **193**, 821
Klein, H. P. & Doudoroff, M. (1950) *J. Bact.* **59**, 739
Klenow, H. & Larsen, B. (1952) *Arch. Biochem. Biophys.* **37**, 488
Kornberg, A. (1950) *J. biol. Chem.* **182**, 805
Kornberg, A. & Pricer, W. E. (1951) *J. biol. Chem.* **193**, 481
Kotel'nikova, A. V. (1948) *C. R. Acad. Sci. U.R.S.S.* **59**, 527
Kubowitz, F. & Ott, P. (1943) *Biochem. Z.* **314**, 94
Kunitz, M. & MacDonald, M. (1946) *J. gen. Physiol.* **29**, 393
Lehmann, H. (1936) *Biochem. Z.* **286**, 336
Leloir, L. F. (1951) In: McElroy, W. D. & Glass, H. B., ed. *Phosphorus metabolism: a symposium on the role of phosphorus in the metabolism of plants and animals*, vol. 1. Johns Hopkins Press, Baltimore
Leloir, L. F., Trucco, R. E., Cardini, C. E., Paladini, A. C. & Caputto, R. (1949) *Arch. Biochem.* **24**, 65
Leuthardt, F. & Testa, E. (1950a) *Helv. chim. Acta*, **33**, 1919
Leuthardt, F. & Testa, E. (1950b) *Helv. physiol. Acta*, **8**, C67
Leuthardt, F. & Testa, E. (1951) *Helv. chim. Acta*, **34**, 931
Lipmann, F., Jones, M. E., Black, S. & Flynn, R. M. (1952) *J. Amer. chem. Soc.* **74**, 2384
Lohmann, K. (1934) *Biochem. Z.* **271**, 264
Lohmann, K. (1935) *Biochem. Z.* **282**, 109
Long, C. (1951) *Biochem. J.* **50**, 407
McElroy, W. D. (1951) *J. biol. Chem.* **191**, 547
Manson, L. A. & Lampen, J. O. (1951) *J. biol. Chem.* **191**, 95
Meyerhof, O. (1927) *Biochem. Z.* **183**, 176
Meyerhof, O. & Green, H. (1950) *J. biol. Chem.* **183**, 377
Meyerhof, O. & Kiessling, W. (1935) *Biochem. Z.* **276**, 239; **280**, 99
Meyerhof, O. & Oesper, P. (1949) *J. biol. Chem.* **179**, 1371
Morton, R. K. (1952) *Animal phosphatases and phosphotransferases*. (Thesis for Ph.D. degree.) University of Cambridge
Muntz, J. A. (1952) *Fed. Proc.* **11**, 263
Najjar, V. A. (1948) *J. biol. Chem.* **175**, 281
Paladinj, A. C., Caputto, R., Leloir, L. F., Trucco, R. E. & Cardini, C. E. (1949) *Arch. Biochem.* **23**, 55
Parnas, J., Ostern, P. & Mann, T. (1934) *Biochem. Z.* **272**, 64
Racker, E. (1947) *J. biol. Chem.* **167**, 843
Rapoport, S. & Luebering, J. (1952) *J. biol. Chem.* **196**, 583
Reid, E., Smith, R. H. & Young, F. G. (1948) *Biochem. J.* **42**, *Proc.* xix
Rowen, J. W. & Kornberg, A. (1951) *J. biol. Chem.* **193**, 497
Sable, H. Z. (1950) *Proc. Soc. exp. Biol., N.Y.* **75**, 215
Sable, H. Z. & Guarino, A. J. (1951) *J. biol. Chem.* **196**, 395
Slein, M. W., Cori, G. T. & Cori, C. F. (1950) *J. biol. Chem.* **186**, 763
Speck, J. H. (1949) *J. biol. Chem.* **179**, 1387; 1405
Stadie, W. C. & Zapp, J. A. (1947) *J. biol. Chem.* **170**, 55
Stern, J. R. & Ochoa, S. (1951) *J. biol. Chem.* **191**, 161
Steyn-Parvé, E. P. (1952) *Biochim. biophys. Acta*, **8**, 310
Stickland, L. H. (1949) *Biochem. J.* **44**, 190
Sutherland, E. W., Cohn, M., Posternak, T. & Cori, C. F. (1949) *J. biol. Chem.* **180**, 1285
Sutherland, E. W., Posternak, T. & Cori, C. F. (1950) *J. biol. Chem.* **179**, 501
Taylor, J. F. (1947) *Fed. Proc.* **6**, 246
Thoai, N. van & Chevillard, L. (1949) *Bull. Soc. Chim. biol., Paris*, **31**, 204
Trucco, R. E., Caputto, R., Leloir, L. F. & Mittelman, N. (1948) *Arch. Biochem.* **18**, 137
Weil-Malherbe, H. (1939) *Biochem. J.* **33**, 1997
Weil-Malherbe, H. & Bone, A. D. (1951a) *Biochem. J.* **49**, 339
Weil-Malherbe, H. & Bone, A. D. (1951b) *Biochem. J.* **49**, 348; 355

* Information available on application to the Ministry of Supply, London.

* * *

SIGNIFICANCE OF BIOCHEMICAL LESIONS IN THE PYRUVATE OXIDASE SYSTEM*

SIR RUDOLPH PETERS M.C. M.D. F.R.C.P. F.R.S.

Department of Biochemistry
University of Oxford

1 Thiamine (vitamin B_1) deficiency
2 Tervalent arsenical poisoning
3 Fluoroacetate poisoning
4 General remarks
References

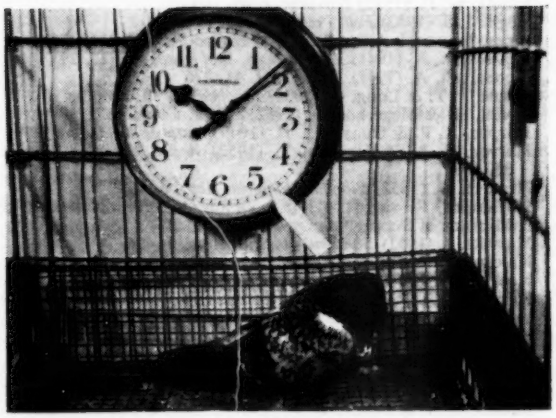
The term " biochemical lesion " (Peters, 1936) was introduced twenty-three years ago, at a time when no proof existed of its reality, to crystallize the idea that pathological disturbances in tissues might be initiated by changes in their biochemistry. If proved, it implies that to rescue medicine from empiricism we should concentrate research upon the initial enzymic changes in a diseased tissue rather than wait upon the developments leading to tissue damage visible to the microscope. Put crudely, it means a substitution of biochemistry for morbid anatomy. Even a decade ago, apart from biochemists, few believed that the enzymic reactions in cells were of much significance for medicine; exceptions were the researches upon cholinesterase—so important for the pharmacology of the nervous system (Dale, 1934)—and those based upon genetical ideas (Garrod, 1923), according to which it was thought that the lack of an enzyme induced by some mutation might be the cause of an " inborn error ". The very fact that biochemical lesions are often reversible means that they provide us with a means of building bridges between purely chemical studies and events in vivo. The term seems useful, as the writer frequently meets it in pathological literature; for recent assessments in relation to clinical and pathological studies, the reader is referred to Bacq (1950), Bacq & Roskam (1951) and Peters (1951). In the present article three biochemical lesions will be considered, which have been studied in Oxford during the last 20 years; detail must perforce be omitted to keep within the space allotted.

1. Thiamine (Vitamin B_1) Deficiency

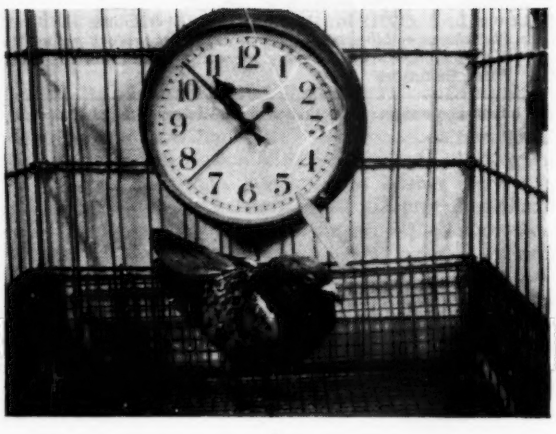
That changes exist which precede the histopathological change can be best illustrated by referring to our old experiments on the biochemical lesion in vitamin B_1 deficiency. The typical opisthotonos and convulsions induced in the pigeon depleted of vitamin B_1 (thiamine, aneurin) have been reversed by the writer within 30 minutes by injecting the vitamin into the subarachnoid space (see fig. 1). This is a central biochemical disturbance, and the speed at which it is cleared by restoration of the vitamin is inconsistent with any view that it could be due to a degenerative change visible with the microscope.

The study of this nutritional deficiency has been fruitful for biochemistry as well as otherwise, because it led to the first definition of the pyruvate oxidase system in brain. The observation was made that a lowered oxygen uptake in vitro with the minced brain from a B_1-deficient animal in a carbohydrate medium was specifically cured by additions of thiamine (Gavrilescu & Peters, 1931); and this was correlated

FIG. 1. CURE OF THIAMINE DEFICIENCY IN PIGEON BY INTRACEREBRAL INJECTION OF THIAMINE



a



b

a : 28 February, 1953, 10.08 a.m. Continuing opisthotonos and rolling convulsions due to rice feeding.
At 10.24 a.m. cerebral injection of 1.5 mg. thiamine-Cl (HCl) + 0.45 mg. glucose in 0.1 ml. saline was given.

b : 10.52 a.m. Showing degree of recovery as indicated by normal stance.

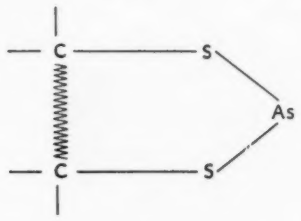
(White arrow indicates time of injection.)

* Based partly on a Noah Morris Lecture delivered in Glasgow in October 1952.

with the degree of the deficiency in vivo (Meiklejohn, Passmore & Peters, 1932). Upon extension of this first observation of the action of a vitamin in vitro, it was found that the deficiency was accompanied by an accumulation of pyruvate ($CH_3.CO.COO^-$) when lactate was the substrate (Peters & Sinclair, 1933; Peters & Thompson, 1934). An increase of the amount of pyruvate in the blood (Thompson & Johnson, 1935) was also found then, and is now a clinical test for thiamine deficiency (Platt & Lu, 1939; Joiner, McArdle & Thompson, 1950, and others); this is one instance where such a test has arisen from pure laboratory research. In the absence of thiamine there is no interference with the metabolic change of glucose or of glycogen to pyruvate; the deficiency interferes selectively with the degradation of pyruvate to carbon dioxide and water by the pyruvate oxidase system; and so it inhibits the liberation of the energy required to maintain the tissue in its living state. The abnormal pathological state is due to withdrawal of a necessary cell constituent. Pyruvic acid itself has not been found toxic, even when injected intravenously (cf. especially Lu & Needham, 1939). Carbohydrate seems uniquely important for brain tissue (Dickens, 1931); evidently the failure in energy supply at the pyruvate step must distort the metabolism of the brain cells and throw the nervous system out of gear. The disastrous effects of the inhibition of this enzyme show that for practical purposes there is no alternative path to the pyruvate one for the main course of metabolism in this tissue.

2. Tervalent Arsenical Poisoning

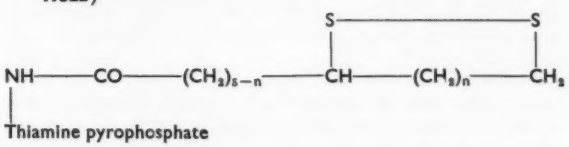
In support of the significance to tissues of an upset of the pyruvate oxidase system, the action of tervalent arsenicals is particularly striking, because these toxic agents make a positive attack, and so eliminate in a different way the coenzyme concerned. It was the feeling that the clue to this poisoning lay in a study of its effects upon the same enzyme system, combined with earlier work on SH substances by ourselves and others, which led the writer to settle down with a group of workers to concentrate upon these studies in 1939, in the hope of finding an antidote. There has appeared previously in *British Medical Bulletin* (Peters, 1948; Thompson, 1948) and elsewhere (Peters, 1952b) an account of the history of the discovery and effects of 2:3-dimercaptopropanol (BAL) upon lewisite blisters, and in systemic poisoning by arsenicals, mercurials and other bivalent metals. Hence it is only necessary here to indicate the relevant points. At close grips the problem developed as follows: though large concentrations of arsenicals attack many enzymes, small amounts (0.015 μM) selectively depress the oxidation of α-keto acids (pyruvate, α-ketoglutarate), as is seen in differential experiments with brain tissue. Like thiamine deficiency, arsenical poisoning increases the amount of pyruvate in blood. Since SH compounds with arsenicals, called thioarsinites, were known to be dissociable (Cohen, King & Strangeways, 1931), excess of the natural tripeptide glutathione should have reversed the toxic effect on the pyruvate oxidation and so freed the enzyme; yet no monothiol was found capable of reversing the action of any tervalent arsenical of lewisite type either in vitro or in vivo. The puzzle was resolved finally by chemical studies of a compound of lewisite with kerateine, a protein product containing SH groups, prepared from hair (Stocken & Thompson, 1946); these led to the idea that the special stability of the thioarsinite formed with some part of the pyruvate oxidase system was due to ring formation between two thiol groups and the arsenical:



The dithiol 2:3-dimercaptopropanol was synthesized; this substance would similarly be expected to form a ring, and might therefore provide a more stable combination with the As fragment than with the enzyme component; both in vitro and in vivo this prediction was fulfilled with a final fortunate dividend in medicine. These researches have been amply confirmed in the USA and elsewhere. As the writer has often urged, it is hardly possible to exaggerate the theoretical significance of this *reversal* of commencing pathological damage (inflammation and oedema) by known chemical action upon a tissue component concerned in the pyruvate oxidation stage of carbohydrate metabolism; it is fundamental for medicine as well as for pathology and pharmacology.

From one angle, the whole problem of the varying effects of arsenicals seems to be reduced to the individual stabilities and chemistry of the thioarsinites formed. Within the last year, another chapter has been added to ours: in the USA recent studies of a bacterial factor, called lipoic acid and concerned with pyruvate oxidation in micro-organisms (Gunsalus, unpublished[1]; Reed & DeBusk, 1952a, b), have led to its isolation as a C_8 carboxylic acid of a dithiol nature with one SH in the ω-position; this is believed to be normally combined with thiamine through the $—NH_2$ of the pyrimidine ring (—CO.NH— linkage) (see fig. 2). This natural dithiol-thiamine coenzyme (lipothiamide) behaves in the bacterial systems to tervalent arsenicals and 2 : 3-dimercaptopropanol as would be predicted from our studies on arsenical poisoning in the animal, and the ring presumably formed with As has the stability to be expected (Whittaker, 1947); i.e., it is rather less stable than the 5-membered ring formed with BAL. If directly applicable to the animal—and this must still be proved—it will provide the final chemical proof that arsenicals

FIG. 2. THE LIPOTHIAMIDE-PYROPHOSPHATE COMPLEX (suggested as coenzyme by Reed & DeBusk, 1952b)

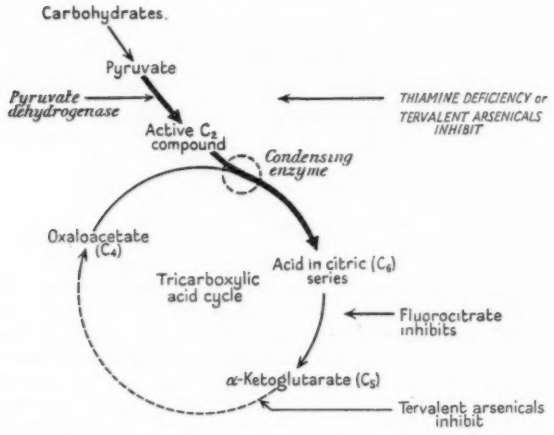


The —COOH of lipoic acid is combined with $—NH_2$ of the pyrimidine ring of thiamine in —CO.NH— linkage.

It seems clear that, in the native pyruvate oxidase system, the SS must exist in SH—SH form to fulfil the phenomena of As toxicity.

[1] Verbal communication at the Second International Congress of Biochemistry held in Paris in July 1952.

FIG. 3. MAIN FEATURES OF PYRUVATE OXIDASE SYSTEM AND SITE OF ACTION OF INHIBITORS



Explanatory notes

1. The total pyruvate oxidase system in brain includes the pyruvate dehydrogenase and the tricarboxylic acid cycle (Krebs, 1943)[2]; probably also coenzyme A.

In absence of a component of the cycle, pyruvate is only degraded to acetate, as indicated originally by Long & Peters (1939), and confirmed by Coxon, Liébecq & Peters (1949) and by others; see the recent work of Schweet & Cheslock, 1952; and of Korkes *et al.* 1951.

2. The pyruvate dehydrogenase is concerned with the initial oxidative decarboxylation to the C_2 fragment; this is represented here as "active C_2 compound". The enzyme requiring Mg consists of apoenzyme and the coenzyme thiamine pyrophosphate plus the "dithiol" (lipothiamide). It is not yet certain whether the "active C_2 compound" is handed on directly from the "dithiol" to the citric-condensing enzyme (Ochoa and colleagues), or whether there is the common intermediate acetyl-S-coenzyme A (Lipmann, Ochoa, Lynen and others).

3. α-Ketoglutarate dehydrogenase is believed to have a coenzyme similar to that of pyruvate; this is not, however, affected in vitamin B_1 deficiency (Peters, 1952d).

4. Acetate (not represented on the diagram) after activation enters the cycle through acetyl-coenzyme A.

5. It should be noted that in high concentration tervalent arsenicals inactivate the condensing enzyme, also succinic dehydrogenase. Di-substituted tervalent arsenicals in low concentration inhibit isolated *iso*citric dehydrogenase (Lotspeich & Peters, 1951), whereas mono-substituted tervalent arsenicals do not inhibit these.

form a ring with SH groups in the pyruvate oxidase system; it will also show clearly that there is a reason for the similarity of the biochemical lesions in thiamine deficiency and arsenical poisoning; they both inactivate by different methods the same entity, the dithiol + thiamine pyrophosphate (? lipothiamide pyrophosphate) necessary for the initial stage in the oxidative decarboxylation of pyruvate to the C_2 fragment and carbon dioxide.

[2] See also Krebs, p. 97 of this Bulletin.

3. Fluoroacetate Poisoning

As it is so dangerous to interfere with the initial stage of pyruvate oxidation, it might be expected that stoppage at a subsequent stage would also matter. Until recently no instance of this was known; the action of the poison fluoroacetate has now provided one (Peters, 1952c). To make its biochemical effect intelligible, some remarks must be devoted to present knowledge of the detailed organization of the pyruvate oxidase system.

Since 1935, when the effect of thiamine was narrowed down by us to an action upon pyruvate, research upon the pyruvate oxidase system in brain and other tissues, by ourselves in Oxford and by others elsewhere, had led to the conclusion that a "battery of enzymes" was needed to cope with this final stage in sugar metabolism (Peters, 1940). Our main facts were confirmed by Sherman & Elvehjem (1936). As Lohmann & Schuster (1937) discovered, the active form of thiamine is the pyrophosphate, a conclusion which was extended to the brain system by Banga, Ochoa & Peters (1939).

The diagram (fig. 3) summarizes the information needed for this review. An interesting recent development, making these researches more concrete, is the theory that the enzymes of the tricarboxylic acid cycle, and so of the total pyruvate oxidase system, are to be found in the mitochondria, that is in microscopically visible components of tissue cells. This present view we owe to the combined efforts of many scientists in Great Britain, the USA and Belgium. The theory has much to support it experimentally, especially now that intact mitochondrial preparations can be prepared for study by fragmenting tissue cells in sucrose solutions (Hogeboom, Schneider & Pallade, 1948). Actually our brain preparations used for many years in the pyruvate work in this laboratory are good preparations of mitochondria (Dawson and Leach, personal communication).

Fluoroacetic acid ($F.CH_2.COOH$) is surprisingly toxic for such a simple substance, whose constitution gives little indication of its poisonous nature. It is especially distinguished by the intense stability of the carbon-fluorine (C-F) bond;

FIG. 4. *DICHAPETALUM CYMOSUM* (HOOK) ENGL. (FAMILY 83. CHAILLETIACEAE)

Burtt Davy (1922), by permission of Journal of Ecology

A : plant, showing habit and underground stems

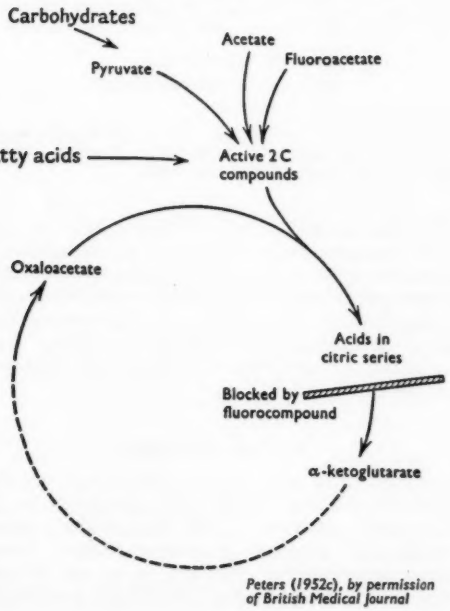
B : flower

C : bifid petal with scale at base

the substance even withstands treatment with hot concentrated sulphuric acid without decomposition. Nevertheless, the C-F bond can evidently be put together in nature, as fluoroacetate is the active principle of the South African poison plant *Dichapetalum cymosum* (fig. 4). Normally in the animal this bond is not broken down. Fluoroacetates are quite non-toxic to isolated enzymes; they are, however, very toxic to the dog, killing at 0.01 mg./kg.; rabbits and guinea-pigs are killed by 0.5 mg./kg., rats by 5 mg./kg. and monkeys by 20–60 mg./kg. (Chenoweth, 1949). In some animals convulsive tetanic signs dominate, and in others (rabbits) effects upon the heart; these always occur after a latent period.

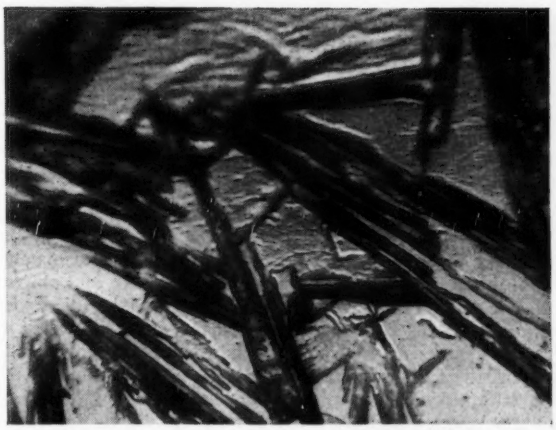
The principle was early enunciated by Saunders and his colleagues (Saunders, 1947) that among a series of C-F compounds (ω-substituted esters of carboxylic acids) of varying length of carbon chain the odd-numbered members were non-toxic, e.g., fluoropropionate, whereas compounds with an even number of carbon atoms were toxic; this could be explained if in each case oxidation proceeded to either a C_2 or a C_3 compound, of which only the former ($F.CH_2CO.$) was then toxic. About 1945, Barron and colleagues (Bartlett & Barron, 1947; Kalnitsky & Barron, 1947) found experimentally that fluoroacetate blocked acetate metabolism in tissue slices, and suggested as a hypothesis that the toxic effect was due to this action of the compound. When Liébecq & Peters (1948, 1949) found that fluoroacetate could lower oxygen uptake in kidney homogenates respiring in fumarate without producing any accumulation of acetate, it became clear that the Barron hypothesis would need some modification; when they also found that citrate accumulated in the fluoroacetate-poisoned preparations, they suggested the hypothesis that fluoroacetate was treated like acetate, and

FIG. 5. DIAGRAM ILLUSTRATING THE "JAMMING" THEORY OF FLUOROACETATE ACTION



Peters (1952c), by permission of British Medical Journal

FIG. 6. CRYSTALS OF FLUOROCITRATE, PREPARED ENZYMICALLY



Peters et al. (1953) by permission of the Royal Society

was synthesized via the tricarboxylic acid cycle condensation to form some compound, which then blocked oxidation by this cycle; this hypothesis was proposed by Liébecq & Peters (1948) independently at about the same time as a similar hypothesis was advanced by Martius (1949), based upon less complete evidence. An important extension of the facts was made when Buffa & Peters (1949) found that animals poisoned with fluoroacetate within one hour or less accumulated large amounts of citrate in most tissues (except liver), there being no simultaneous increase of α-ketoglutarate; so it became possible to formulate the "jamming" hypothesis as in the diagram (see fig. 5).

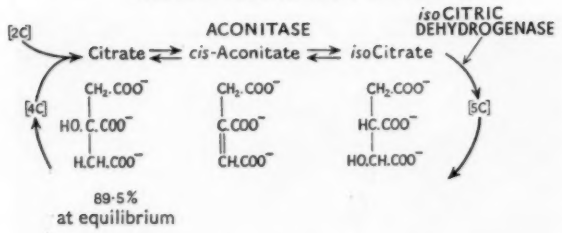
It should be mentioned that the accumulation of citrate under these conditions is the best evidence we have that the tricarboxylic acid cycle functions in vivo, and that the salient facts have been confirmed and extended by others (Potter & Busch, 1950; Lindenbaum, White & Schubert, 1951).

It has now been proved (Peters, Wakelin & Buffa, 1952; Peters *et al.* 1953) that a synthesis of the fluoroacetyl fragment has taken place by the isolation from kidney homogenates poisoned with fluoroacetate of small amounts of a crystalline preparation of a monofluorotricarboxylic acid (fig. 6), which has been identified spectrochemically by comparison of its infra-red spectrum with that of synthetic fluorocitric acid, prepared by D. E. A. Rivett. The presence of fluorine was established by Gillieson & Newcombe (1952), using a spectrochemical method. It can be inferred from our earlier work that the effect of the inhibitory fluoro factor is selective at the citrate stage.

The fact that citric acid accumulates in the animal poisoned with fluoroacetate suggested strongly that fluorocitrate blocked either the enzyme aconitase or the *iso*citric dehydrogenase (see fig. 7).

Research has shown (Lotspeich, Peters & Wilson, 1952; Peters & Wilson, 1952) that the isolated aconitase is blocked by fluorocitrate and that, at any rate in part, the blocking is competitive. It is significant that in a mitochondrial preparation from kidney the inhibitory action of the fluoro*citrate* on citrate oxidation is much more potent than upon the isolated

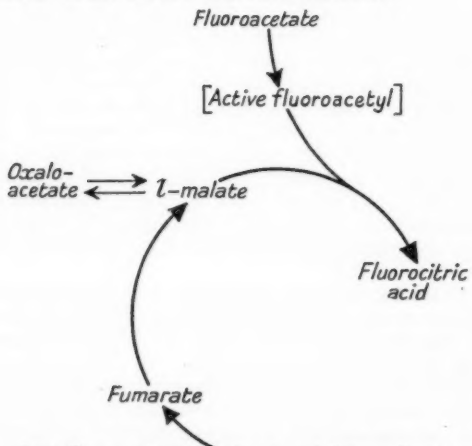
FIG. 7. TRICARBOXYLIC ACID REACTIONS OF TRICARBOXYLIC ACID CYCLE



aconitase in solution; as little as 0.6 μg. (approximately 0.003 μmole) suffices to inhibit the disappearance of 1 μmole of citric acid. The above facts appear to prove that the toxic action of fluoroacetate is due to what the writer has called a "lethal synthesis", by which the fluoroacetate is synthesized enzymically to fluorocitric acid (fig. 8); the "jamming" of the tricarboxylic acid cycle occurs at the aconitase stage, and the effect is magnified by organization within the mitochondrion. This lethal synthesis is the opposite of the usual protective synthesis and has other scientific implications (Peters, 1952a).

A final proof of the lethal nature of fluorocitrate was wanted. It was thought that this could be obtained, if fluorocitrate could be proved to be toxic under conditions where fluoroacetate was non-toxic. Such conditions hold in vitro with pigeon-brain tissue, as, unlike kidney tissue, it does not form the inhibitor from fluoroacetate. Figure 9 shows that the addition of fluorocitrate to particles from pigeon brain respiring in pyruvate lowers oxygen uptake and increases the accumulation of citrate, in agreement with a blocked aconitase. This work (Peters & Wakelin, 1953), together with that of Coxon, Liébecq & Peters (1949) and Coxon & Peters (1950), in the writer's opinion finally concludes the argument that

FIG. 8. DIAGRAMMATIC SUMMARY OF SOME RECENT EXPERIMENTAL FACTS REGARDING THE SYNTHESIS OF FLUORINE-CONTAINING INHIBITOR FROM FLUOROACETATE

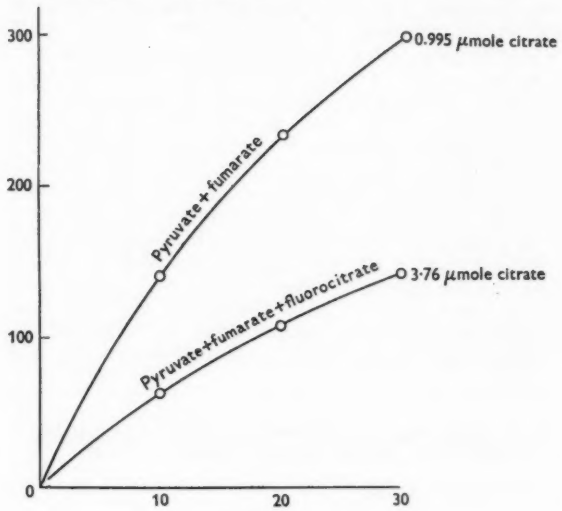


Recent experimental facts (see Peters, 1952c) suggest that fluoroacetate synthesis may not follow exactly the same path as acetate.

the pyruvate oxidase system in brain includes the tricarboxylic acid cycle.

We may now consider what happens in vivo with pigeon brain. Quite recently (June, 1952), the writer injected into the subarachnoid space of some pigeons small quantities (50–80 μg.) of fluorocitrate. The result was astonishing. The animal recovered from the anaesthetic (ether) within about 4 minutes. It remained reasonably normal for 10–20 minutes after which a period of agitation commenced, terminating in violent convulsions somewhat like those seen in vitamin B_1 deficiency, though often there is more wing movement (see

FIG. 9. THE EFFECT OF CITRATE INHIBITOR (FLUOROCITRATE) ON THE OXYGEN UPTAKE AND CITRATE ACCUMULATION IN HOMOGENATES OF PIGEON BRAIN



Abscissae : time (min.)
Ordinates : oxygen uptake (μl.)
Substrate : pyruvate + fumarate

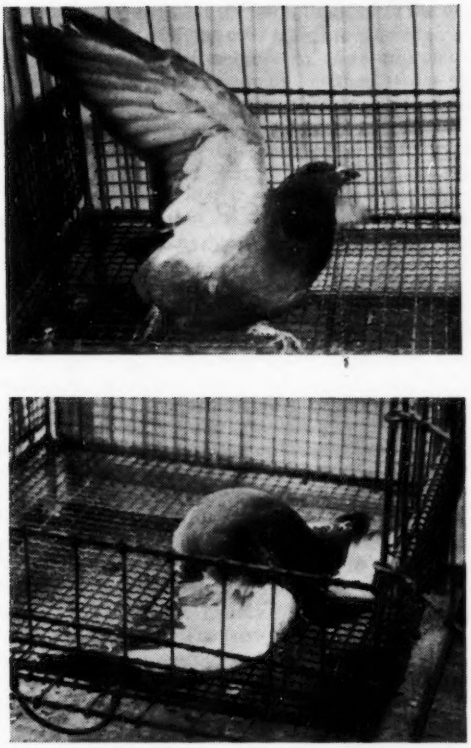
As shown by the lower curve, the addition of fluorocitrate lowers oxygen uptake by stopping the metabolism of citrate, which therefore accumulates (Peters & Wakelin, 1953).

fig. 10). The condition always terminates in death, and there is more citrate present at death than in the normal. Saline injections are without effect, as well as fluoroacetate injections even when given in amounts larger than would be the equivalent of the fluorocitrate. This clinical state finally proves the lethal nature of the synthesis in fluoroacetate poisoning. It appears to be induced by "jamming" the pyruvate oxidase system in vivo at the citrate stage.

4. General Remarks

In all these three cases, there is interference at some stage of the pyruvate oxidase system and convulsions are induced in the pigeon. (We have recently also obtained convulsions by intracranial injection of lewisite (Peters, Wakelin & Hastings, unpublished).) They can be set side by side with the view that oxygen poisoning is induced by interference with

FIG. 10. PIGEON IN CONVULSIONS AFTER INJECTION OF FLUOROCITRATE INTO SUBARACHNOID SPACE



The upper photograph shows a stage of active wing movement, and the lower one a stage of convulsion with head retraction.

pyruvate oxidation (Dickens, 1946; Mann & Quastel, 1946); and that in epidemic dropsy pyruvate also accumulates (Wilson & Ghosh, 1937; Sarkar, 1948). It is reasonable to suppose that the primary event is interference with the energy supply to the cell; especially does this seem likely since it is known that interference with glucose metabolism by failure of the oxygen supply, or by cutting off glucose whether by starvation or overdoses of insulin, produces similar effects (Peters, 1930). But when this has been said, there still remains unsolved the problem of the steps by which the failure of the energy supply induces the convulsive state. It is well known, especially from the studies of Hastings and his colleagues (Mullin, Hastings & Lees, 1938), that alterations in the ratio of inorganic ions in the cerebrospinal fluids rapidly induce hyperexcitability. In confirmation of this we have found that the injection of citrate (2–3 μmoles) into the subarachnoid space of the pigeon immediately induces a convulsion (Hastings, Peters & Wakelin, unpublished), which can be relieved equally rapidly by the injection of calcium ions. For the moment, however, no bridge has been built between these two sets of phenomena. All attempts to cure the convulsions of fluorocitrate poisoning with calcium injections, by others (Eeg-Larsen & Naess, 1951) as well as by ourselves, have failed; and it can be mentioned that, so far, injections of calcium and magnesium salts have not saved animals injected with fluorocitrate (Hastings, Peters & Wakelin, unpublished results). Nevertheless, reflection based upon the modern idea that these enzymes are enclosed in the mitochondria suggests that reversal of fluoroacetate poisoning will be difficult to accomplish; it implies that any therapeutic agent must pass the two permeability barriers of the cell surface and of the mitochondrion inside the cell. However satisfied we may feel with the limited progress towards logical therapy made by studying these biochemical lesions, we must realize that there are still stiff problems ahead; these are less likely to be solved by studies in vitro of artificial mixtures of biochemical substances than by experiments relating in vitro to in vivo conditions.

REFERENCES

Bacq, Z. M. (1950) *Principes de physiopathologie et de thérapeutique générales*. Masson, Paris

Bacq, Z. M. & Roskam, J. (1951) In: *XXVIII^e Congrès Français de Médecine, Bruxelles, 1951. Rapports*, vol. 1, p. 212. Charles Lorie, Brussels; (1951) *Pr. méd.* **59,** 1483 [Abstract]

Banga, I., Ochoa, S. & Peters, R. A. (1939) *Biochem. J.* **33,** 1109

Bartlett, G. R. & Barron, E. S. G. (1947) *J. biol. Chem.* **170,** 67

Buffa, P. & Peters, R. A. (1949) *J. Physiol.* **110,** 488

Chenoweth, M. B. (1949) *J. Pharmacol.* **1,** 383

Cohen, A., King, H. & Strangeways, W. I. (1931) *J. chem. Soc.*, p. 3043

Coxon, R. V., Liébecq, C. & Peters, R. A. (1949) *Biochem. J.* **45,** 320

Coxon, R. V. & Peters, R. A. (1950) *Biochem. J.* **46,** 300

Dale, H. H. (1934) *Proc. R. Soc. Med.* **28,** 319

Davy, J. Burtt (1922) *J. Ecol.* **10,** 211

Dickens, F. (1931) *Biochem. J.* **30,** 661

Dickens, F. (1946) *Biochem. J.* **40,** 145

Eeg-Larsen, N. & Naess, K. (1951) *Acta pharm. tox., Kbh.* **7,** 133

Garrod, A. E. (1923) *Inborn errors of metabolism*, 2nd ed. Oxford Medical Publications, Oxford

Gavrilescu, N. & Peters, R. A. (1931) *Biochem. J.* **25,** 1397; 2150

Gillieson, A. H. C. P. & Newcombe, R. A. (1952) *Biochem. J.* **50,** *Proc.* xiv

Hogeboom, G. H., Schneider, W. C. & Pallade, G. E. (1948) *J. biol. Chem.* **172,** 619

Joiner, C. L., McArdle, B. & Thompson, R. H. S. (1950) *Brain* **73,** 431

Kalnitsky, G. & Barron, E. S. G. (1947) *J. biol. Chem.* **170,** 83

Korkes, S., Del Campillo, A., Gunsalus, I. C. & Ochoa, S. (1951) *J. biol. Chem.* **193,** 721

Krebs, H. A. (1943) *Advanc. Enzymol.* **3,** 191

Liébecq, C. & Peters, R. A. (1948) *J. Physiol.* **108,** 11*P*

Liébecq, C. & Peters, R. A. (1949) *Biochim. biophys. Acta*, **3,** 215

Lindenbaum, W., White, M. R. & Schubert, J. (1951) *J. biol. Chem.* **190,** 585

Lohmann, K. & Schuster, P. (1937) *Biochem. Z.* **294,** 188

Long, C. & Peters, R. A. (1939) *Biochem. J.* **33,** 759

Lotspeich, W. D. & Peters, R. A. (1951) *Biochem. J.* **49,** 704

Lotspeich, W. D., Peters, R. A. & Wilson, T. H. (1952) *Biochem. J.* **51,** 20

Lu, G. D. & Needham, D. M. (1939) *Biochem. J.* **33,** 1544

Mann, P. J. G. & Quastel, J. H. (1946) *Biochem. J.* **40,** 139

Martius, C. (1949) *Liebigs Ann.* **561,** 227

Meiklejohn, A. P., Passmore, R. & Peters, R. A. (1932) *Proc. roy. Soc.* B, **111,** 391

Mullin, F. J., Hastings, A. B. & Lees, W. M. (1938) *Amer. J. Physiol.* **121,** 719

Peters, R. A. (1930) *J. State Med.* **38,** 3

Peters, R. A. (1936) *Lancet*, **1,** 1161

Peters, R. A. (1940) *Chem. & Ind. (Rev.)* **59,** 373

Peters, R. A. (1948) *Brit. med. Bull.* **5,** 313

Peters, R. A. (1951) *Edinb. med. J.* **58,** 197

Continued at foot of page 122

FOLIC ACID AND RELATED COMPOUNDS IN THE METABOLISM OF MICRO-ORGANISMS

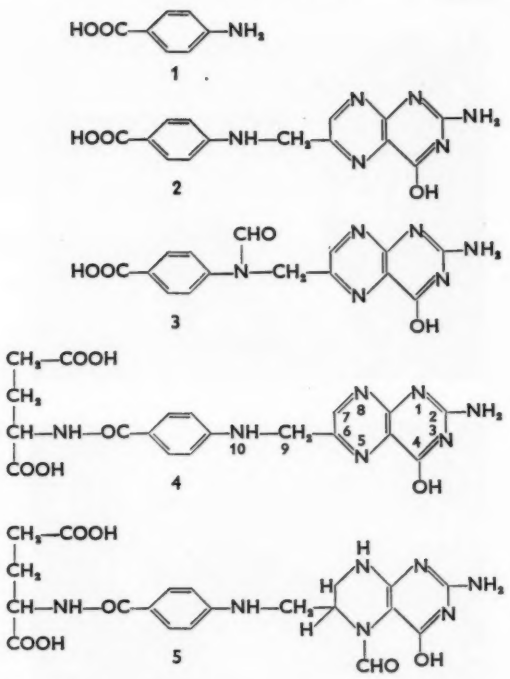
D. D. WOODS M.A. Ph.D. F.R.S.

Microbiology Unit, Department of Biochemistry
University of Oxford

1 Metabolic interrelationship of folic acid and *p*-aminobenzoic acid (*p*-AB)
2 Replacement of *p*-aminobenzoic acid (*p*-AB) and folic acid for the growth of micro-organisms
3 Function of *p*-aminobenzoic acid (*p*-AB) and folic acid at defined steps in biosynthetic pathways
a Localization of effect
b Discovery of an intermediate in purine synthesis
c Work with cell suspensions of micro-organisms
d Active form of folic acid
e General nature of the reactions in which folic acid is concerned
f Mechanism of function of folic acid
4 Conclusions
References

The subject to be discussed illustrates a quite different approach, based on research on the nutrition of micro-organisms, to problems of enzyme action; it is of special application to enzyme systems concerned in the synthetic mechanisms of the cell. Folic acid and related substances (fig. 1), generally classified as members of the vitamin B group, are essential growth factors for various micro-organisms and were first characterized chemically as such; they are also essential for the normal metabolism and growth of many other creatures, including birds and mammals. Those micro-organisms which do not require an external source appear to be able to synthesize these substances for themselves. The first member of the group shown to have biological activity was *p*-aminobenzoic acid (*p*-AB); it was discovered not through any direct effect on bacterial growth, but through its ability to overcome inhibition by the sulphonamide drugs

FIG. 1. STRUCTURE OF VARIOUS MEMBERS OF THE FOLIC ACID GROUP



1: p-aminobenzoic acid
2: pteroic acid
3: rhizopterin
4: pteroylglutamic acid
5: probable structure of folinic acid (N^5-formyltetrahydropteroylglutamic acid)

(Woods, 1940). The hypothesis then put forward that *p*-AB was an essential metabolite, and that the sulphonamides act (by virtue of their analogous chemical structure) by competing for the enzyme concerned in the utilization of *p*-AB, has been amply confirmed, particularly by the identification of *p*-AB as an actual growth factor for many micro-organisms; this

BIOCHEMICAL LESIONS IN THE PYRUVATE OXIDASE SYSTEM

Continued from page 121

Peters, R. A. (1952a) *Proc. roy. Soc.* B, **139,** 143
Peters, R. A. (1952b) *J. R. Inst. publ. Hlth,* **15,** 89; 122; 151
Peters, R. A. (1952c) *Brit. med. J.* **2,** 1165
Peters, R. A. (1952d) *II^e Congrès International de Biochimie, Paris, 21–27 Juillet 1952.* In: *Symposium sur le cycle tricarboxylique,* p. 69. Société d'Edition d'Enseignement Supérieur, Paris
Peters, R. A. & Sinclair, H. M. (1933) *Biochem. J.* **27,** 1910
Peters, R. A. & Thompson, R. H. S. (1934) *Biochem. J.* **28,** 916
Peters, R. A. & Wakelin, R. W. (1953) *J. Physiol.* **119,** 421
Peters, R. A., Wakelin, R. W. & Buffa, P. (1952) *Biochem. J.* **50,** *Proc.* xiii
Peters, R. A., Wakelin, R. W., Buffa, P. & Thomas, L. C. (1953) *Proc. roy. Soc.* B, **140,** 497
Peters, R. A. & Wilson, T. H. (1952) *Biochim. biophys. Acta,* **9,** 310
Platt, B. S. & Lu, G. D. (1939) *Biochem. J.* **33,** 1525
Potter, V. R. & Busch, H. (1950) *Cancer Res.* **10,** 353
Reed, L. J. & DeBusk, B. G. (1952a) *J. Amer. chem. Soc.* **74,** 3457; 3964
Reed, L. J. & DeBusk, B. G. (1952b) *J. biol. Chem.* **199,** 881
Sarkar, S. N. (1948) *Nature, Lond.* **162,** 265
Saunders, B. C. (1947) *Nature, Lond.* **160,** 179
Schweet, R. S. & Cheslock, K. (1952) *J. biol. Chem.* **199,** 749
Sherman, W. C. & Elvehjem, C. A. (1936) *Amer. J. Physiol.* **117,** 142
Stocken, L. A. & Thompson, R. H. S. (1946) *Biochem. J.* **40,** 529
Thompson, R. H. S. (1948) *Brit. med. Bull.* **5,** 319
Thompson, R. H. S. & Johnson, R. E. (1935) *Biochem. J.* **29,** 694
Whittaker, V. P. (1947) *Biochem. J.* **41,** 56
Wilson, H. E. C. & Ghosh, B. K. (1937) *Indian med. Gaz.* **72,** 147

hypothesis is now generally accepted. The actual enzyme has not yet been isolated nor has folic acid so far been found to be a component of any isolated enzyme system.

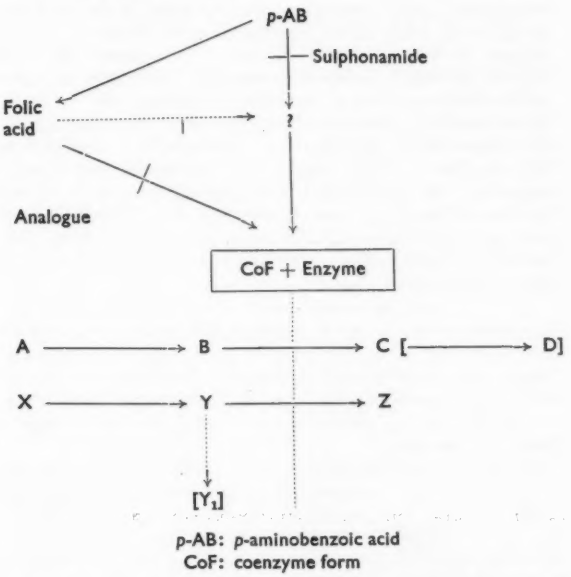
The common requirement for the substances of the vitamin B group by all types of living cells points to a basic similarity in their general pattern of metabolism. Any information as to the function of these vitamins obtained from micro-organisms is likely therefore to be of wider application. In all cases in which the function of these factors has been precisely determined, they have proved to be part of the structure of coenzymes or prosthetic groups of key enzyme systems, e.g., nicotinic acid in the phosphopyridine nucleotide coenzymes of respiration, riboflavin in the flavoprotein enzymes, vitamin B_6 (as pyridoxal) in the pyridoxal phosphate coenzymes of amino-acid metabolism, pantothenate in coenzyme A required for the transfer of two-carbon residues, thiamine in the cocarboxylase of pyruvate metabolism. It is reasonable to adopt a similar idea as a working hypothesis in the case of vitamin B-group substances of unknown function. If the enzyme system concerned is required for some synthetic process of the cell, valuable initial information as to the general nature of the reaction catalysed may often be obtained from experiments with growing cultures of micro-organisms, where the rate of synthesis of new cell material is so high.

The case of folic acid may be taken as an example of this method of approach and its subsequent development. It will be necessary to restrict the discussion mainly to work with micro-organisms, and, since the results to be summarized cover (either directly or by implication) over a hundred papers, reference must be made mainly to reviews where detailed citations will be found. The term "folic acid" will be used in general reference to the group as a whole; individual members have been used mostly as synthetic materials (for terminology see legend to fig. 1).

1. Metabolic Interrelationship of Folic Acid and *p*-Aminobenzoic Acid (*p*-AB)

Folic acid contains a *p*-AB residue in the molecule; the existence of a series of bacteria requiring one or other member of the group as minimal requirement for growth—e.g., *Clostridium acetobutylicum* (requiring *p*-AB), *Streptococcus faecalis* R (rhizopterin, pteroic acid), *Lactobacillus casei* (pteroylglutamic acid), *Leuconostoc citrovorum* (citrovorum factor, folinic acid)—would suggest that these compounds represent intermediates of increasing complexity in the biosynthesis of the ultimate coenzyme form of folic acid. An analogous series is found in the case of nicotinic acid (see Knight, 1945). However, the inability of the higher forms to replace either the growth factor or antisulphonamide function of *p*-AB with some organisms (though they do with others) is against this simple view. Nevertheless, cell suspensions of representative organisms of these various types all synthesize a folic acid from *p*-AB and the synthesis is inhibited by sulphonamides in a way which parallels their effect on growth (Lascelles & Woods, 1952). It remains possible, though rather unlikely, that with some organisms (though not others) *p*-AB has a second function in cell metabolism not exerted through its intermediate conversion to folic acid. The simplest interpretation of the facts at present known (Woods, 1952a, 1952b; but see also Shive, 1952) is that neither rhizopterin, pteroylglutamic acid, nor folinic acid are direct and successive intermediates between *p*-AB and the hypothetical coenzyme

FIG. 2. SCHEME ILLUSTRATING THE INTERPRETATION OF CERTAIN TYPES OF NUTRITIONAL EXPERIMENTS



p-AB is *p*-aminobenzoic acid and CoF its coenzyme form. C, D and Z are essential metabolites which become growth factors if the organism cannot synthesize them; A, B, X and Y are intermediates in their synthesis by competent organisms. Y_1 indicates a by-product which may be formed from Y.

form (CoF), but are converted to such intermediates with varying efficiency by some organisms and not at all by others; *Leuconostoc mesenteroides*, for example, responds to *p*-AB and to folinic acid but not to pteroylglutamate. Unless properties of differing permeability to various members of the folic acid group by closely related organisms are invoked, it seems unlikely that the most complex known form (folinic acid) can itself be CoF.

2. Replacement of *p*-Aminobenzoic Acid (*p*-AB) and Folic Acid for the Growth of Micro-organisms

The ability of mixtures of certain other substances (other than folic acid) to replace *p*-AB either for growth or in overcoming sulphonamide inhibition gave the first clue as to the ultimate function of the folic acid group. The principles involved are best made clear by a generalized schematic representation (fig. 2). Suppose that *p*-AB gives rise to some form of folic acid (a reaction inhibited competitively by sulphonamides) which in turn is converted to the coenzyme form (CoF). Analogues of folic acid (e.g. aminopterin) prevent the growth of organisms requiring folic acid and presumably also inhibit some step in the formation of CoF. Suppose that the coenzyme is part of an enzyme system required for the conversion of B to C and of X to Y, in two synthetic pathways leading to formation of cell constituents. If the formation of CoF is prevented, growth will fail. This may be done either (i) by omitting *p*-AB or folic acid from the growth medium of organisms requiring these vitamins, or

(ii), in the case of organisms able to synthesize them, by adding sulphonamides or folic acid analogues. It would be expected that growth would occur under both these conditions if the substances C and Z were provided preformed in the medium; furthermore, such growth should now be insensitive to inhibition by sulphonamides. If only C is provided, then the amount of *p*-AB required for growth or to overcome sulphonamide inhibition would be reduced, though not to zero. When therefore it is found that a substance (or combination of substances) of chemical type different from the growth factor is able to replace that growth factor, we have a plain suggestion that the factor is concerned at some stage of the synthesis of the compounds in question. The technique using inhibitory analogues has been much developed by Shive (see reviews by Shive, 1950, 1951, 1952) and has been called by him " inhibition analysis."

It is clear that the simple interpretations stated above may not always be valid and that other explanations may exist in some cases for the replacement of growth factors or for the antagonism of inhibitory analogues. This kind of evidence should be considered a sign-post to further work rather than as a proof that the growth factor takes part in the synthetic process in question.

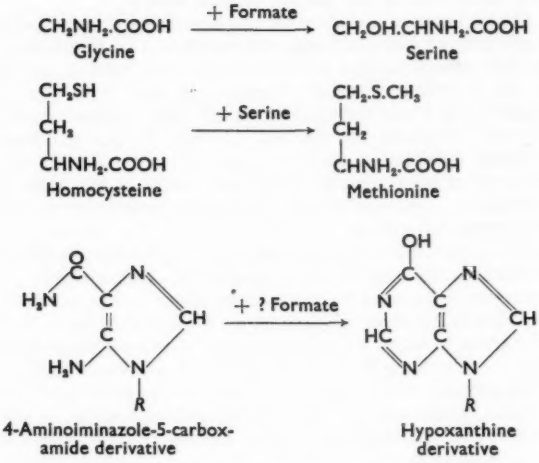
Research along these lines points to the following substances as being products resulting ultimately from the function of *p*-AB and folic acid in various micro-organisms (for details see reviews by Shive, 1950, 1951; Woods, 1950, 1952a, 1952b; Woolley, 1952): (i) the pyrimidine, thymine, or more probably its deoxyriboside, thymidine, (ii) purines, the precise derivative varying with the organism tested, (iii) methionine and other amino-acids, notably serine, histidine, leucine, lysine and (iv) vitamin B_{12}, which contains a benziminazole derivative (an analogue of purine) in the molecule. On the present hypothesis, therefore, CoF would be part of an enzyme system catalysing some step in the synthesis of each of these substances. Since the latter are many and varied, the question arises as to whether CoF is concerned in a number of different reactions involving the same type of chemical change (cf. coenzyme A in the transfer of two-carbon residues) or in the synthesis of some common precursor. Further progress towards localizing the effect of CoF thus depends on knowledge of the main intermediates in each synthetic pathway; here again, studies of nutrition have proved most helpful.

3. Function of *p*-Aminobenzoic Acid (*p*-AB) and Folic Acid at Defined Steps in Biosynthetic Pathways

a. Localization of Effect

When some of the main intermediates are already known, further " replacement " experiments may show the step at which the growth factor is concerned. Thus, in the scheme of fig. 2, if C and D are equally effective in replacing *p*-AB, then the latter presumably functions at some step before C; if, however, D is active whilst C is not, then it is probable that CoF functions in the conversion of C to D. In the case of methionine synthesis there is strong evidence from the nutritional requirements of biochemical mutants of micro-organisms (see review by Horowitz, 1950) that cysteine, cystathionine and homocysteine are successive intermediates. The inability of homocysteine to replace the *p*-AB-sparing effect of methionine suggests a function of CoF at this stage (see fig. 3). Similar and other evidence (see Woods, 1952a, 1952b) points to the conversion of glycine to serine as the CoF-catalysed step in the synthesis of the latter amino-acid.

FIG. 3. REACTIONS IN WHICH FOLIC ACID PROBABLY HAS A FUNCTION



b. Discovery of an Intermediate in Purine Synthesis

When growth is impeded by partial lack of *p*-AB and a reaction such as $Y \rightarrow Z$ (see fig. 2) is restricted, then Y (or a by-product such as Y_1 formed from it) may continue to be formed and accumulate in the medium. The isolation of 4-aminoiminazole-5-carboxamide (fig. 3) from cultures of *Bacterium coli* partially inhibited by sulphonamide (see Shive, 1950) provided the first clue that this substance (or a derivative) was an intermediate in purine synthesis and that *p*-AB and ultimately CoF were required for its conversion to purine. It has also now been isolated from cultures of the same organism containing the folic acid analogue aminopterin, and from purine-requiring and *p*-AB-requiring mutants (see Woolley, 1952; Gots & Chu, 1952). The biological properties of the carboxamide make it improbable that the free substance is itself an intermediate[1]; furthermore there is good evidence from work with animal tissues (Schulman & Buchanan, 1952) that condensation with ribosephosphate occurs before the addition of a carbon atom and ring closure. These matters are discussed more fully by Shive (1951, 1952), Gots & Chu (1952) and Woods (1952a).

c. Work with Cell Suspensions of Micro-organisms

A more rigid definition of the substrates of the enzyme systems depending on CoF requires experiments with systems simpler than whole growing cultures in which synthesis of the product is inferred from the general increase in cell substance, and in which other components of the growth medium (or metabolic products) besides those under test may also be substrates; any synthesis is likely to be at least a bimolecular reaction. The use of washed suspensions of cells

[1] The accumulation of a riboside of the carboxamide in young cultures of *Bact. coli* in the presence of sulphadiazine has now been shown (Greenberg, 1952). (Author's note added in proof.)

specifically deficient in the growth factor is the most useful technique for following up the suggestions provided by the growth experiments (if synthesis does indeed occur under these conditions with appropriate additions of defined substrates and growth factor). Cells deficient in either *p*-AB or folic acid may be obtained by growth in a medium containing either (i) suboptimal quantities of the growth factor, or (ii) no growth factor but the appropriate mixture of the substances which replace it. The second method is to be preferred since it results in a good yield of cells which are totally deficient in these growth factors. The use of vitamin-deficient animal tissues is of course a well-tried method for the study of metabolic function; it was first used by Peters & Thompson (1934) in their work on the function of vitamin B_1 in pigeon brain.[2]

Some recent work of the Oxford Microbiology Unit has been devoted to applying these methods to the bacterial syntheses of serine and methionine (see review by Woods, 1952a). Cell suspensions of *Str. faecalis* R deficient in folic acid, and of *Leuconostoc mesenteroides* and *Saccharomyces cerevisiae* deficient in *p*-AB, were all found to synthesize serine when incubated in a simple mixture containing glycine, formate, glucose and buffer, but only if the appropriate growth factors were also added; vitamin B_6 was also required. In similar work with cells of a mutant of *Bact. coli* requiring *p*-AB, synthesis of methionine was obtained from buffered homocysteine plus glucose but again only if *p*-AB was added. The other reactant, i.e. the source of the one-carbon residue for the methylation, appears to be serine, since cells of another mutant which cannot synthesize serine were unable to form methionine under these conditions unless serine was also added. The first mutant still required *p*-AB if serine was supplied. This work clearly defines the substrates required for the two syntheses; glucose metabolism is no doubt an ancillary process yielding energy for the reaction.

d. Active Form of Folic Acid

It was pointed out earlier that, so far as can be judged from growth experiments, none of the forms of folic acid so far known has the properties expected for the functional form (CoF). In the work with cell suspensions and serine synthesis (see Woods, 1952a) the form required corresponded with the form active for growth (viz., pteroylglutamate or folinic acid for *Str. faecalis*, *p*-AB or folinic acid for *Leuconostoc mesenteroides*), except that, with *S. cerevisiae*, folinic acid as well as *p*-AB was active for synthesis though not for growth. Some evidence was obtained with *Str. faecalis* that pteroylglutamate must be converted to a formylated derivative for activity. It is clear that research on the nature of the coenzyme form would be greatly aided by the isolation and purification of one of the enzymes concerned; this has not yet been achieved.

[2] See Peters, p. 116 of this Bulletin.

e. General Nature of the Reactions in which Folic Acid is Concerned

In all three cases where the function of folic acid has been traced to a specific step in an over-all synthesis it has proved to be a reaction in which a single-carbon unit is added, viz., glycine to serine, homocysteine to methionine, 4-aminoiminazole-5-carboxamide derivative to hypoxanthine derivative (see fig. 3). Indeed, when Shive first isolated the carboxamide, he proposed that *p*-AB or a derivative acted as coenzyme for the addition of a carbon atom and ring closure (Shive, 1950). Such a function for CoF (transfer of single-carbon residues) is now a generally accepted hypothesis which may be used in attempting to elucidate the precise nature of the other reactions in which folic acid appears to be concerned. Since no pyrimidine derivative other than thymine or thymidine has been described as having any action in sparing the requirements of *p*-AB or folic acid for growth (or antagonizing sulphonamides or folic acid analogues), it seems likely that CoF is concerned here with the addition of the methyl group (possibly before ring closure) rather than with a single-carbon incorporation during the formation of the pyrimidine ring. A step at which a one-carbon residue is added during the synthesis of histidine is indicated by the work of Levy & Coon (1951), who found incorporation by yeast of labelled formate into this amino-acid. There is as yet no definite evidence that folic acid intervenes at these particular steps in the synthesis of thymine and histidine.

f. Mechanism of Function of Folic Acid

If CoF transfers single-carbon residues, it is an attractive hypothesis that it does so by the intermediate addition and removal of the —CHO group of its own molecule; folinic acid, for example, has such a group (fig. 1). However, Broquist *et al.* (1951) have found tetrahydropteroylglutamate to have much less biological activity than its formyl derivative (folinic acid). Moreover, Elwyn, Weissbach & Sprinson (1951), in isotope experiments with animal tissues (incorporation of the β-carbon of serine into choline and thymine), find evidence against this view.

4. Conclusions

It is probable that folic acid has an ultimate coenzyme-like function in cell metabolism in a number of reactions in which single-carbon residues are transferred. The precise chemical nature of the coenzyme form and the mechanism of its action are not yet determined.

REFERENCES

Broquist, H. P., Fahrenbach, M. J., Brockman, J. A., Stokstad, E. L. R. & Jukes, T. H. (1951) *J. Amer. chem. Soc.* **73**, 3535
Elwyn, D., Weissbach, A. & Sprinson, D. B. (1951) *J. Amer. chem. Soc.* **73**, 5510
Gots, J. S. & Chu, E. C. (1952) *J. Bact.* **64**, 537
Greenberg, G. R. (1952) *J. Amer. chem. Soc.* **74**, 6307
Horowitz, N. H. (1950) *Advanc. Genet.* **3**, 33
Knight, B. C. J. G. (1945) *Vitam. & Horm.* **3**, 105
Lascelles, J. & Woods, D. D. (1952) *Brit. J. exp. Path.* **33**, 288
Levy, L. & Coon, M. J. (1951) *J. biol. Chem.* **192**, 807
Peters, R. A. & Thompson, R. H. S. (1934) *Biochem. J.* **28**, 916
Schulman, M. P. & Buchanan, J. M. (1952) *J. biol. Chem.* **196**, 513
Shive, W. W. (1950) *Ann. N.Y. Acad. Sci.* **52**, 1212
Shive, W. W. (1951) *Vitam. & Horm.* **9**, 75
Shive, W. W. (1952) *Annu. Rev. Microbiol.* **6**, 437
Woods, D. D. (1940) *Brit. J. exp. Path.* **21**, 74
Woods, D. D. (1950) *Ann. N.Y. Acad. Sci.* **52**, 1199
Woods, D. D. (1952a) *II^e Congrès International de Biochimie, Paris, 21–27 Juillet 1952.* In : *Symposium sur le metabolisme microbien*, p. 86. Société d'Edition d'Enseignement Supérieur, Paris
Woods, D. D. (1952b) *Bull. World Hlth Org.* **6**, 35
Woolley, D. W. (1952) *A study of antimetabolites.* Chapman & Hall, London

β-GLUCURONIDASE AND RELATED ENZYMES

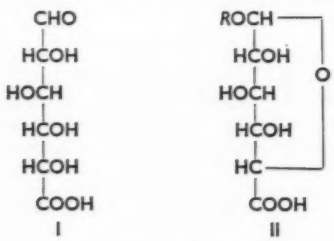
G. A. LEVVY D.Sc. F.R.I.C.

Rowett Research Institute
Bucksburn, Aberdeenshire

1 Assay
2 Purification and kinetics
3 Specificity
4 Inhibitors
5 Tissue factors influencing β-glucuronidase activity
6 β-Glucuronidase and glucuronide synthesis
7 β-Glucuronidase and hyaluronidase
8 Changes in β-glucuronidase activity in vivo
References

Interest in β-glucuronidase has been stimulated by the suggestions that it may be involved in steroid metabolism (Fishman, 1947) and in tissue proliferation (Levvy, Kerr & Campbell, 1948), and by its use as a reagent for the smooth hydrolysis of steroid glucuronides (see, for example, Bayliss, 1952). β-Glucuronidase assay has been suggested as a clinical test in toxaemia of pregnancy (Odell & McDonald, 1948) and in uterine carcinoma (Odell & Burt, 1950; see also Fishman, Kasdon & Homburger, 1950).

Glucuronic acid (I) is widely distributed in living tissues. In the animal body, many alcohols and phenols, including certain steroid hormones and stilboestrol (4 : 4′-dihydroxy-αβ-diethylstilbene), combine with glucuronic acid to form β-D-glucuronides (general formula, II; R = aglycone residue) and are excreted as such in the urine. Glucuronic acid also occurs in animal tissues in large molecules such as hyaluronic acid, chondroitin and heparin. Certain bacterial capsular mucopolysaccharides have been shown to contain glucuronic acid, and it is present in some of the ill-defined polysaccharides that constitute hemicellulose, and in many of the plant gums.



I II

Decomposition of a urinary glucuronide by an animal tissue preparation was first noted by Röhmann (1908), but the enzyme responsible was not identified until 1934 (Masamune, 1934). This enzyme, β-glucuronidase, catalyses the hydrolysis of β-D-glucuronides to the aglycones and glucuronic acid (Masamune, 1934; Levvy, 1948), and it has been found in almost every animal tissue studied.

Other enzymes which liberate the aglycones from β-D-glucuronides have been demonstrated in various species of bacteria (Barber, Brooksbank & Haslewood, 1948; Buehler, Katzman & Doisy, 1949; Smith & Mills, 1950; Barber, Brooksbank & Kuper, 1951), in certain wood-rotting fungi (Marsh, Oxford & Levvy, unpublished), in the medium-sized micro-organisms in the rumen of the sheep (Karunairatnam & Levvy, 1951), in locust crop liquor (Robinson, Smith & Williams, 1952), in the digestive tract of marine molluscs (Dodgson, Lewis & Spencer, 1952), and in the contents of the small and large intestines from many species of animal (Marsh, Alexander & Levvy, 1952). Pending studies of their specificity and mode of action, these enzymes can only provisionally be classified as β-glucuronidases. Their specificities may differ widely from that of the enzyme in animal tissues, and they may liberate the aglycone from a glucuronide by some means other than simple hydrolysis, for example, through the combined actions of a decarboxylase and a xylosidase. In the present paper, the term β-glucuronidase is reserved for the enzyme in animal tissues, and the others are referred to as glucuronide-decomposing enzymes.

The roots of *Scutellaria baicalensis* contain an enzyme, baicalinase, that decomposes baicalin, the glucuronide of 5:6:7-trihydroxyflavone, but not menthol β-D-glucuronide (Miwa, 1935). The relationship of this enzyme to those discussed above is obscure.

1. Assay

Biosynthetic glucuronides isolated from urine have been used exclusively as substrates in the assay of β-glucuronidase and other glucuronide-decomposing enzymes, with the exception of baicalinase. Methods in which the enzyme activity is measured by determining the reducing power of the glucuronic acid liberated from menthol glucuronide (Masamune, 1934; Fishman, 1939a) have given way to others in which the free aglycone is determined colorimetrically. For most purposes, the best method of assay is that in which phenolphthalein glucuronide is used as substrate (Talalay, Fishman & Huggins, 1946). The procedure is simple, calling for little, if any, purification of the enzyme preparation, and the colour reaction given by the free aglycone with alkali is sensitive and specific.

2. Purification and Kinetics

Following earlier work by Fishman (1939a) and Graham (1946), three quite distinct methods have recently been published for the purification of β-glucuronidase (Sarkar & Sumner, 1950; Bernfeld & Fishman, 1950a; Smith, 1951; Smith & Mills, 1953). All three apparently result in about 1,000-fold purification on a dry-weight or nitrogen basis. The product, a globulin, is activated by deoxyribonucleic acid (Bernfeld & Fishman, 1950b). No comparable studies of the purification of other glucuronide-decomposing enzymes have yet appeared.

Mills (1948) discovered that the enzyme from ox spleen comprised two, or more (Mills & Paul, 1949; Mills, Paul & Smith, 1953a), fractions, differing slightly in the optimum pH for hydrolysis of a glucuronide (see also Kerr, Campbell & Levvy, 1949). Several other workers have, however, failed to detect more than one fraction in the enzyme from this and other tissues (Sarkar & Sumner, 1950; Bernfeld & Fishman, 1950a; Morrow, Carroll & Greenspan, 1951; Rossiter & Wong, 1950). Bernfeld & Fishman (1950c) observed that the

pH-activity curve for β-glucuronidase is altered in shape by the addition of adenylic, deoxyribonucleic, and other acids (see also Fishman, 1952). This effect is also produced by suramin and phthalic acid (Smith & Mills, 1953), and by the anionic detergent, Teepol XL (Levvy & Marsh, unpublished). Varying the substrate concentration causes remarkable changes in the pH-activity curve, particularly at pH values below 4 and above 6 (Mills *et al.* 1953a). In accordance with the suggestion of Bernfeld & Fishman (1950c), it seems plausible to suppose that, in any one species at least, the enzyme is a single protein which can form complexes with various compounds before or after preparation of the tissue extract, and which can thus display different pH optima, depending upon the tissue of origin, the purification procedure, and the final test medium. Mills *et al.* (1953a), however, consider that each of their fractions is a distinct protein.

Table I gives pH optima for β-glucuronidase and some other glucuronide-decomposing enzymes. Most β-glucuronidase preparations cause maximal hydrolysis of phenolphthalein glucuronide at one or other of the pH values shown,

TABLE I. pH OPTIMA FOR β-GLUCURONIDASE AND SOME OTHER GLUCURONIDE-DECOMPOSING ENZYMES

Source	Substrate	pH Optimum	Reference
Animal tissues	Phenolphthalein glucuronide	4.5 and 5.2	Mills (1948)
Bacterium coli	Phenolphthalein glucuronide	6.2	Buehler *et al.* (1949)
Rumen liquor	Phenolphthalein glucuronide	6.1*	Karunairatnam & Levvy (1951)
Locust crop liquor	8-Hydroxyquinoline glucuronide	4.5	Robinson *et al.* (1952)

* Varies slightly with the purity of the preparation (Marsh, unpublished).

and replacing this substrate by others leads to little change in the optimum or optima. Mills *et al.* (1953a) and Smith & Mills (1953) report additional optima at pH 3.4 and pH 6.3, but these were seen only at certain substrate concentrations, and have still to be confirmed. Table II shows values for K_m, the dissociation constant of the active enzyme-substrate complex, for β-glucuronidase and various substrates. The reciprocal of K_m gives the affinity of the enzyme for a compound, and Table II also shows the relative affinities for

TABLE II. DISSOCIATION CONSTANTS (K_m) OF THE ACTIVE ENZYME-SUBSTRATE COMPLEX FOR β-GLUCURONIDASE AND VARIOUS β-D-GLUCURONIDES

(Affinity for phenolphthalein glucuronide taken as unity)

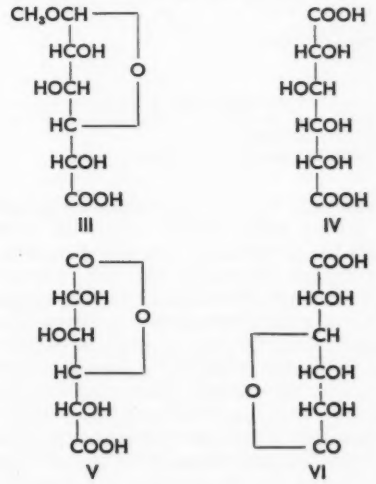
Aglycone	K_m (M)	Relative affinity	Reference
Phenolphthalein	0.00005	1.0	Talalay *et al.* (1946)
Oestriol . .	0.0005	0.1	Fishman (1939b)
Phenol . .	0.0035	0.014	Kerr *et al.* (1948)
Menthol . .	0.004	0.013	Fishman (1939b)
Borneol . .	0.01	0.005	Fishman (1939b)
p-Chlorophenol .	0.03	0.0017	Spencer & Williams (1951)

the different substrates. It should be noted that inhibition of β-glucuronidase by excess substrate sometimes leads to difficulties in determining K_m (cf. Kerr, Graham & Levvy, 1948).

3. Specificity

Masamune (1934) found that β-glucuronidase would hydrolyse neither menthol α-D-glucuronide nor various α- and β-D-glucosides (see also Graham, 1946). The recent development of a convenient method for the synthesis of hexuronides by oxidation of the corresponding hexosides (Marsh, 1951, 1952; Barker, Bourne & Stacey, 1951) has facilitated further studies of the specificity. Hexuronides tested with mouse-liver β-glucuronidase (Levvy & Marsh, 1952) were menthol α-D-glucuronide, methyl α- and β-D-galacturonides and mannuronides, and α- and β-D-glucuronic acid-1-phosphate (II; $R = H_2PO_3$). At the same time various glycosides were tested, as were also methyl α- and β-D-glucofuranuronides (Owen, Peat & Jones, 1941), supplied by Professor F. Smith, Department of Agriculture, University of Minnesota, USA. Only β-glucuronic acid-1-phosphate was hydrolysed by β-glucuronidase. Discrimination was made between the actions of β-glucuronidase and phosphatase by the use of saccharo-1:4-lactone (see section 4). It would appear that β-glucuronidase is specific, not just for β-glucuronides, but for β-glucopyranuronides, since methyl β-D-glucofuranuronide (III) was not attacked. The synthesis of *laevo*-menthol β-D-glucuronide (Marsh, 1951, 1952), identical with the natural product, affords final proof that this compound and therefore others hydrolysed by β-glucuronidase have the pyranose structure shown in II.

Details of the specificities of other glucuronide-decomposing enzymes have not been published, but work in progress (Marsh, unpublished) suggests that the enzyme from the sheep's rumen acts only on β-glucuronides.



4. Inhibitors

Karunairatnam & Levvy (1949) discovered that saccharic (glucosaccharic) acid (IV) is a powerful inhibitor for β-glucuronidase, whilst mucic (galactosaccharic) acid is much weaker. Saccharic acid is remarkably non-toxic to animals

Carr, 1947; Karunairatnam & Levvy, 1949). Seeking an inhibitor which might produce toxic effects in vivo as a result of its action on β-glucuronidase, Levvy (1952) prepared 3-methylsaccharo-1:4-lactone (Levene & Meyer, 1922), and from it 3-methylsaccharate. On test in vitro, only the lactone inhibited the enzyme. In the light of these results, the action of saccharic acid was reinvestigated, and it was found that saccharate as such is not an inhibitor. The active agent is saccharo-1:4-lactone (V), formed in small amounts in saccharate solutions at the slightly acid pH of enzyme assay. The inhibitory action of mucate solutions also is apparently due to lactone formation.

Except in ring size, saccharo-1:4-lactone (V) closely resembles a β-D-glucuronide (II) in structure. Saccharo-3:6-lactone (VI) has little, if any, effect on β-glucuronidase. Lactonization is increased by heating a saccharate or mucate solution, and when a solution of potassium hydrogen saccharate is boiled, one-third is converted into the 1:4-lactone after 30 minutes. This is a convenient way of making an inhibitor solution, since saccharo-1:4-lactone is difficult to obtain pure (Smith, 1944).

Table III gives data for the inhibitors discussed, all of which act competitively. Saccharo-1:4-lactone is a very powerful inhibitor, with an affinity for β-glucuronidase 240

TABLE III. INHIBITORS OF β-GLUCURONIDASE (Levvy, 1952)

(Affinity of phenolphthalein glucuronide taken as unity)

Inhibitor	Concentration for 50% inhibition (10^{-4} M)*	Apparent relative affinity†
Saccharate solution (freshly prepared)	1,800	0.8
Saccharate solution (boiled for 30 min.)	12	81.0
Saccharo-1:4-lactone	3.6	240.0
3-Methylsaccharo-1:4-lactone	1,600	1.0
Mucate solution (freshly prepared)‡	—	0.02
Mucate solution (boiled for 2 hr.)	725	1.5

* Substrate 0.00063 M-phenolphthalein glucuronide, pH 5.2.

† Calculated from values for K_i, the dissociation constant of the enzyme-inhibitor complex.

‡ $5{,}000 \times 10^{-6}$ M caused 6% inhibition.

times greater than that of phenolphthalein glucuronide, the substrate with the highest known affinity (cf. Table II). The term affinity has of course no theoretical significance in the case of saccharate or mucate solutions, but the apparent values give a measure of the amounts of active agent present.

Inhibition of β-glucuronidase by saccharate solutions has been repeatedly confirmed, and has been employed to prevent hydrolysis of oestriol glucuronide in urine from women in labour (Clayton & Marrian, 1950), and to terminate the reaction in the enzyme assay (Robinson *et al.* 1952). Campbell (1949) found that saccharate prevented the hydrolysis of 8-hydroxyquinoline glucuronide in a histochemical method for the localization of β-glucuronidase (Friedenwald & Becker, 1948), and this observation has been utilized in a critical study of different histochemical tests for this enzyme (Burton & Pearse, 1952). All other glucuronide-decomposing enzymes tested were inhibited by "saccharate" solutions or by saccharo-1:4-lactone (Smith & Mills, 1950; Robinson *et al.* 1952; Levvy, 1952; Marsh *et al.* 1952). Unfortunately, neither saccharo-1:4-lactone nor 3-methylsaccharo-1:4-lactone is stable under physiological conditions.

With the exception of glucuronic acid (Spencer & Williams, 1951), all the other compounds found to inhibit β-glucuronidase do so only in relatively high concentrations (about 10^{-2} M), and the enzyme is unaffected by many of the common enzyme poisons, such as fluoride, iodoacetate, cyanide and azide; the status of ascorbic and citric acids as inhibitors is obscure (see Wong & Rossiter, 1951; Mills *et al.* 1953a). β-Glucuronidase is sometimes inhibited by an unknown substance in animal tissues (see section 5).

5. Tissue Factors Influencing β-Glucuronidase Activity

The phenolphthalein glucuronide method made it possible to assay β-glucuronidase directly in homogenates of animal tissues in water. In earlier methods of assay, insoluble cell debris had been discarded, usually after a short period of incubation in acetate or citrate buffer. The distribution of enzyme activity in tissue homogenates has been studied in a series of papers (Kerr & Levvy, 1951; Walker & Levvy, 1951, 1953; Walker, 1952) which will be considered together. From the results it appears that considerable hazards attend the comparative measurement of tissue β-glucuronidase activity, and still more so attempts to deduce the physiological function of the enzyme from such measurements.

In water homogenates of mouse liver, 40% of the total enzyme activity in adult liver, and a somewhat smaller fraction of the activity in infant liver, was found in the cytoplasmic granules. In dilute preparations at least, this fraction of the enzyme was fully active. It was spread over granules of all sizes, so that sedimentation was complete only at a relatively high centrifugal force (25,000 *g*). Buffering preparations to a slightly acid pH, however, led to agglutination and ready sedimentation of the granules. The enzyme was released, without appreciable change in activity, from the granules by mechanical disruption, by addition of a surface-active agent, or by prolonged incubation in acetate, but not citrate, buffer. When mouse liver was homogenized in isotonic sucrose or sodium chloride solution, in order to preserve the subcellular particles, nearly all the enzyme was found in these bodies. In this type of preparation, as compared with a water homogenate, the enzyme did not display full activity, and appeared to be rather inaccessible to substrate. Dilution with water caused osmotic swelling of the particles, and partial release of enzyme from them all, irrespective of size; the residual enzyme became more accessible to substrate, and reached full activity when the enzyme released approached the proportions in a water homogenate. Activation was irreversible.

Rat liver resembled mouse liver in the distribution of the enzyme in different types of homogenate. There was, however, an important difference between the two tissues in that rat-liver β-glucuronidase was reversibly inhibited by an unidentified, thermostable and non-dialysable tissue constituent, apparently similar to the blood plasma anti-glucuronidase of Fishman, Altman & Springer (1948). Inhibition, which varied inversely with the age of the animal, was overcome by the surface-active agent, Triton X-100. The type of inhibition was unusual, the inhibitor combining with the enzyme-substrate complex, so that its effect varied directly with the substrate concentration. Mouse liver contained a similar inhibitor for rat-liver β-glucuronidase, although the enzyme in mouse liver was immune to its action.

6. β-Glucuronidase and Glucuronide Synthesis

This has been the subject of a recent discussion (Levvy & Fishman, 1952), to which the reader is referred for further details. Fishman (1940, 1947) considers that β-glucuronidase is responsible for the synthesis of glucuronides in the animal body. Only Florkin *et al.* (1942) have been successful in obtaining evidence for the condensation in vitro of an aglycone with glucuronic acid in the presence of β-glucuronidase. Condensation was very slight and the authors did not consider that their results had any bearing on the problem of glucuronide synthesis in vivo. Glucuronide synthesis is readily demonstrated in surviving liver slices (Lipschitz & Bueding, 1939; Crépy, 1947; Levvy & Storey, 1949; Karunairatnam, Kerr & Levvy, 1949; Storey, 1950), in which it is depressed by inhibitors of oxidation and of phosphorylation, but not by saccharo-1:4-lactone; it is not stimulated by glucuronic acid. Unlike β-glucuronidase, the glucuronide-synthesizing system is confined to liver and kidney. Changes in liver β-glucuronidase activity in vivo (see section 8) are not paralleled in the synthetic system.

On theoretical grounds, Storey (1950) suggested that the immediate precursor in the synthesis of glucuronides in vivo is glucuronic acid-1-phosphate, but Levvy & Marsh (1952) and Touster & Reynolds (1952) could obtain no evidence for this in experiments with liver slices. If β-glucuronidase could be shown to act as a transferase, it would reconcile Fishman's view, that this enzyme controls the conjugation of alcohols and phenols with glucuronic acid, with the evidence that synthesis is ultimately due to quite a different enzyme system. There is as yet no hint that β-glucuronidase can act in this way.

Radioactive tracer techniques have recently been applied to this problem in whole animal experiments (Mosbach & King, 1950; Packham & Butler, 1952; Eisenberg & Gurin, 1952; Douglas & King, 1952). The results, although not entirely convincing, suggest that glucose or glucuronic acid can be incorporated in the glucuronide molecule without breaking the 6-carbon chain, but of course provide no information on the mechanism of conjugation.

7. β-Glucuronidase and Hyaluronidase

Hyaluronic acid is a mucopolysaccharide from animal tissues, in which the repeating unit is a glucuronide of *N*-acetylglucosamine (Meyer, 1947). Crude preparations of testicular hyaluronidase, an enzyme that degrades both hyaluronic acid and chondroitin sulphate (Meyer & Rapport, 1950), have recently been shown to contain β-glucuronidase (Meyer, Linker & Rapport, 1951; see also Hahn, 1946, 1947). β-Glucuronidase hydrolyses the oligosaccharides that are the end-products of the action of purified testicular hyaluronidase on hyaluronic acid (Meyer *et al.* 1951); this incidentally indicates that the repeating unit has a β-glucopyranuronide structure (see section 3). Although Meyer *et al.* (1951) consider it unlikely that the breakdown of hyaluronic acid can proceed thus far in the animal body, it is interesting to note that β-glucuronidase could be a common link in the metabolism of glucuronidogenic steroid hormones and of mucopolysaccharides.

8. Changes in β-Glucuronidase Activity in Vivo

Most of the work in this field was done in ignorance of the effects of the tissue factors described in section 5, and diverse procedures, not always completely specified, have been employed in preparing the enzyme for assay. Comparison of results obtained by different workers is thus not always possible, and there is at least one instance in which the use of an inappropriate procedure has yielded results diametrically opposed to the truth (Walker & Levvy, 1953). Nevertheless, some of the changes in β-glucuronidase activity that have been observed in vivo are too well substantiated to be dismissed as artifacts.

Fishman (1940) discovered that administration of menthol to mice and borneol to dogs caused a rise in the β-glucuronidase activity of liver, kidney and spleen, but not of uterus and other sex organs. Since both of these compounds form glucuronides, Fishman thought that this was an adaptive response by the enzyme, acting synthetically, to the presence of excess aglycone. His theory was extended to cover the increase in the enzymic activity of uterus, as distinct from organs like liver, that followed the treatment of ovariectomized mice with oestrogens (Fishman & Fishman, 1944; Fishman, 1947). The additional assumption was made that in its postulated synthetic role the enzyme differed in its specificity for different aglycones according to its site of action in the body. No such specificity was, however, observed in the properties of the different enzyme preparations in vitro.

Levvy and his collaborators (Levvy *et al.* 1948; Kerr *et al.* 1949; Kerr, Campbell & Levvy, 1950), whilst confirming Fishman's principal findings, showed that variation in β-glucuronidase activity was a much wider phenomenon than he had imagined. Anything that caused liver or kidney damage in mice produced a rise in the β-glucuronidase activity of the tissue, and the action of menthol was found to come into this category. The rise in enzymic activity seemed to be associated with the later phase of tissue repair rather than with the actual injury. Progesterone and testosterone antagonized the effects of oestrogens on the enzyme in uterus, and in infant mice the enzyme activity of any organ was higher than in adults. Elevated tissue β-glucuronidase activities could be reduced to the normal adult values by the administration of colchicine. On the basis of these findings, it was suggested that the β-glucuronidase activity of a tissue was related in some unknown fashion to its state of proliferation.

This theory readily explains the high β-glucuronidase activity of human cancer tissue (Fishman & Anlyan, 1947; Odell & Burt, 1949; see, however, Mills & Smith, 1951), and the variation in the activity of the enzyme in human endometrium during the menstrual cycle (Odell & Fishman, 1950). On the other hand, the theory is not entirely satisfactory in explaining results obtained in a study of peripheral nerve undergoing Wallerian degeneration (Hollinger *et al.* 1952), and in detailed investigations of the effects of sex hormones in mice and rats (Harris & Cohen, 1951; Knobil, 1952). The most serious criticism of the theory comes from the work of Mills, Paul & Smith (1953b), who found that liver β-glucuronidase activity was lower in new-born than in adult rats, a finding which has been confirmed by Walker & Levvy (1953) in experiments in which the effects of the endogenous inhibitor on rat-liver β-glucuronidase (see section 5) were excluded. Complete understanding of the changes in activity which β-glucuronidase undergoes in vivo must await elucidation of the function of the enzyme by work in other directions, for example, by the use of a selective inhibitor or by the study of its possible role in mucopolysaccharide metabolism.

References

Barber, M., Brooksbank, B. W. L. & Haslewood, G. A. D. (1948) *Nature, Lond.* **162,** 701
Barber, M., Brooksbank, B. W. L. & Kuper, S. W. A. (1951) *J. Path. Bact.* **63,** 57
Barker, S. A., Bourne, E. J. & Stacey, M. (1951) *Chem. & Ind.* **45,** 970
Bayliss, R. I. S. (1952) *Biochem. J.* **52,** 63
Bernfeld, P. & Fishman, W. H. (1950a) *Arch. Biochem.* **27,** 475
Bernfeld, P. & Fishman, W. H. (1950b) *Science,* **112,** 653
Bernfeld, P. & Fishman, W. H. (1950c) *Fed. Proc.* **9,** 150
Buehler, H. J., Katzman, P. A. & Doisy, E. A. (1949) *Fed. Proc.* **8,** 189
Burton, J. F. & Pearse, A. G. E. (1952) *Brit. J. exp. Path.* **33,** 87
Campbell, J. G. (1949) *Brit. J. exp. Path.* **30,** 548
Carr, C. J. (1947) *Proc. Soc. exp. Biol., N.Y.* **65,** 189
Clayton, B. E. & Marrian, G. F. (1950) *J. Endocrinol.* **6,** 332
Crépy, O. (1947) *Arch. Sci. physiol.* **1,** 427
Dodgson, K. S., Lewis, J. I. M. & Spencer, B. (1952) *Biochem. J.* **51,** *Proc.* xlii
Douglas, J. F. & King, C. G. (1952) *J. biol. Chem.* **198,** 187
Eisenberg, F. & Gurin, S. (1952) *J. biol. Chem.* **195,** 317
Fishman, W. H. (1939a) *J. biol. Chem.* **127,** 367
Fishman, W. H. (1939b) *J. biol. Chem.* **131,** 225
Fishman, W. H. (1940) *J. biol. Chem.* **136,** 229
Fishman, W. H. (1947) *J. biol. Chem.* **159,** 7
Fishman, W. H. (1952) In: *Ciba Foundation Colloquia on Endocrinology.* Vol. I: *Steroid hormones and tumour growth, and steroid hormones and enzymes,* p. 231. Churchill, London
Fishman, W. H., Altman, K. I. & Springer, B. (1948) *Fed. Proc.* **7,** 154
Fishman, W. H. & Anlyan, A. J. (1947) *Cancer Res.* **7,** 808
Fishman, W. H. & Fishman, L. W. (1944) *J. biol. Chem.* **152,** 487
Fishman, W. H., Kasdon, S. C. & Homburger, F. (1950) *J. Amer. med. Ass.* **143,** 350
Florkin, M., Crismer, R., Duchateau, G. & Houet, R. (1942) *Enzymologia,* **10,** 220
Friedenwald, J. S. & Becker, B. (1948) *J. cell. comp. Physiol.* **31,** 303
Graham, A. F. (1946) *Biochem. J.* **40,** 603
Hahn, L. (1946) *Ark. Kemi Min. Geol.* **21,** 1
Hahn, L. (1947) *Biochem. Z.* **318,** 123
Harris, R. S. & Cohen, S. L. (1951) *Endocrinology,* **48,** 264
Hollinger, D. M., Logan, J. E., Mannell, W. A. & Rossiter, R. J. (1952) *Nature, Lond.* **169,** 670
Karunairatnam, M. C., Kerr, L. M. H. & Levvy, G. A. (1949) *Biochem. J.* **45,** 496
Karunairatnam, M. C. & Levvy, G. A. (1949) *Biochem. J.* **44,** 599
Karunairatnam, M. C. & Levvy, G. A. (1951) *Biochem. J.* **49,** 210
Kerr, L. M. H., Campbell, J. G. & Levvy, G. A. (1949) *Biochem. J.* **44,** 487
Kerr, L. M. H., Campbell, J. G. & Levvy, G. A. (1950) *Biochem. J.* **46,** 278
Kerr, L. M. H., Graham, A. F. & Levvy, G. A. (1948) *Biochem. J.* **42,** 191
Kerr, L. M. H. & Levvy, G. A. (1951) *Biochem. J.* **48,** 209
Knobil, E. (1952) *Endocrinology,* **50,** 16
Levene, P. A. & Meyer, G. M. (1922) *J. biol. Chem.* **54,** 805
Levvy, G. A. (1948) *Biochem. J.* **42,** 2
Levvy, G. A. (1952) *Biochem. J.* **52,** 464
Levvy, G. A. & Fishman, W. H. (1952) *Science,* **116,** 285
Levvy, G. A., Kerr, L. M. H. & Campbell, J. G. (1948) *Biochem. J.* **42,** 462
Levvy, G. A. & Marsh, C. A. (1952) *Biochem. J.* **52,** 690
Levvy, G. A. & Storey, I. D. E. (1949) *Biochem. J.* **44,** 295
Lipschitz, W. L. & Bueding, E. (1939) *J. biol. Chem.* **129,** 333
Marsh, C. A. (1951) *Nature, Lond.* **168,** 602
Marsh, C. A. (1952) *J. chem. Soc.,* p. 1578
Marsh, C. A., Alexander, F. & Levvy, G. A. (1952) *Nature, Lond.* **170,** 163
Masamune, H. (1934) *J. Biochem., Tokyo,* **19,** 353
Meyer, K. (1947) *Physiol. Rev.* **27,** 335
Meyer, K., Linker, A. & Rapport, M. M. (1951) *J. biol. Chem.* **192,** 275
Meyer, K. & Rapport, M. M. (1950) *Arch. Biochem.* **27,** 287
Mills, G. T. (1948) *Biochem. J.* **43,** 125
Mills, G. T. & Paul, J. (1949) *Biochem. J.* **44,** *Proc.* xxxiv
Mills, G. T., Paul, J. & Smith, E. E. B. (1953a) *Biochem. J.* **53,** 232
Mills, G. T., Paul, J. & Smith, E. E. B. (1953b) *Biochem. J.* **53,** 245
Mills, G. T. & Smith, E. E. B. (1951) *Science,* **114,** 690
Miwa, T. (1935) *Acta phytochim., Tokyo,* **8,** 231
Morrow, A. G., Carroll, D. M. & Greenspan, E. M. (1951) *J. nat. Cancer Inst.* **11,** 663
Mosbach, E. H. & King, C. G. (1950) *J. biol. Chem.* **185,** 491
Odell, L. D. & Burt, J. C. (1949) *Cancer Res.* **9,** 362
Odell, L. D. & Burt, J. C. (1950) *J. Amer. med. Ass.* **142,** 226
Odell, L. D. & Fishman, W. H. (1950) *Amer. J. Obstet. Gynec.* **59,** 200
Odell, L. D. & McDonald, D. F. (1948) *Amer. J. Obstet. Gynec.* **56,** 74
Owen, L. N., Peat, S. & Jones, W. J. G. (1941) *J. Chem. Soc.,* p. 339
Packham, M. A. & Butler, G. C. (1952) *J. biol. Chem.* **194,** 349
Robinson, D., Smith, J. N. & Williams, R. T. (1952) *Biochem. J.* **50,** *Proc.* xiii
Röhmann, F. (1908) *Biochemie.* Springer, Berlin
Rossiter, R. J. & Wong, E. (1950) *Canad. J. Res.* (E), **28,** 69
Sarkar, N. K. & Sumner, J. B. (1950) *Arch. Biochem.* **27,** 453
Smith, E. E. B. (1951) *Biochem. J.* **50,** *Proc.* viii
Smith, E. E. B. & Mills, G. T. (1950) *Biochem. J.* **47,** *Proc.* xlix
Smith, E. E. B. & Mills, G. T. (1953) *Biochem. J.* **54,** 164
Smith, F. (1944) *J. chem. Soc.,* p. 633
Spencer, B. & Williams, R. T. (1951) *Biochem. J.* **48,** 537
Storey, I. D. E. (1950) *Biochem. J.* **47,** 212
Talalay, P., Fishman, W. H. & Huggins, C. (1946) *J. biol. Chem.* **166,** 757
Touster, O. & Reynolds, V. H. (1952) *J. biol. Chem.* **197,** 863
Walker, P. G. (1952) *Biochem. J.* **51,** 223
Walker, P. G. & Levvy, G. A. (1951) *Biochem. J.* **49,** 620
Walker, P. G. & Levvy, G. A. (1953) *Biochem. J.* **54,** 56
Wong, E. & Rossiter, R. J. (1951) *Canad. J. med. Sci.* **29,** 195

* * *

THE FORMATION OF PEPTIDES IN ENZYMIC REACTIONS

CHARLES S. HANES Ph.D. Sc.D. F.R.S.

Department of Biochemistry, University of Toronto, Canada

1 Biological processes of condensation
2 Synthesis of peptide bonds in ATP-dependent condensations
3 Transpeptidation and transamidation reactions
a Transfers of α-amino-acyl groups
b Transfers of γ-glutamyl groups
4 Discussion
References

THE search for enzymic mechanisms concerned in the synthesis of protein has been stimulated by the progress made in recent years toward identifying the reactions involved in the biosynthesis of other complex molecules which resemble proteins in being " condensation polymers ", that is, formed from smaller structural units joined by condensations with the elimination of the elements of water. Such condensed molecules, large and small, form major constituents of living matter; they display much diversity in the chemical nature of their structural units, and also accordingly in the nature of the anhydro-bonds through which these units are linked together.

1. Biological Processes of Condensation

A traditional hypothesis has been that the condensed state might be attained in the cell by direct reversal of the hydrolytic process, and accordingly in reactions catalysed by the same enzymes that catalyse hydrolysis. Since in a majority of cases equilibrium between a condensed molecule and its split products lies far on the side of hydrolysis, appreciable synthesis by direct condensations between the split products could occur only if the condensed molecules were being removed effectively from the system. The possibility that such mechanisms may come into play in special cases is not excluded—and indeed a speculative suggestion regarding such a case will be included later—but it is clear that the bulk of the condensations leading to the elaboration of condensed cellular constituents are indirect in nature.

The study of such processes began with the classical observations of Harden & Young (1905) on the formation of hexosephosphates accompanying the fermentation of sugars in cell-free preparations of " zymase " from yeast. The analysis of this complex system, and of comparable systems from other sources, notably muscle tissue, has attracted the attention of many distinguished investigators during the intervening half-century. One of the consequences of these fundamental studies has been the recognition of the types of enzymic reaction involved in the synthesis of some variety of naturally occurring condensed molecules. Two distinct classes of reaction comprise such processes of synthesis and these may be termed the primary condensation reactions, and the group transfer reactions.

The primary condensation reactions, probably relatively few in number, lead to the synthesis de novo of condensed molecules; it is the anhydro-linkages of these products of primary condensation which, as a result of being substituted and re-substituted, supply the condensing capacity of the cell. While knowledge of this class of reaction as a whole remains imperfect, the known examples are the so-called coupled oxidation and phosphorylation reactions. In these, carboxyl phosphates (or related condensed molecules) are formed, the result apparently of dehydrogenation of a carbonyl addition compound. The formation of a carboxyl phosphate or similar bond is accompanied by a large increase in the free energy, as was shown by Lipmann (1941), so that an essential condition for its formation in significant amounts is that there be a concomitant transformation characterized by a compensating decrease in the free energy, e.g., the oxidation.

The other class of reactions known to form stages in the synthesis of different types of condensed molecule (phosphoric esters, simple and complex saccharides, certain dinucleotides, etc.) are the group transfer reactions. First amongst these, and possibly linking all synthetic pathways to the primary condensation reactions, stands the formation of adenosinetriphosphate (ATP) by transfer to adenosinediphosphate (ADP) of a phosphoryl group from a carboxyl phosphate.

A group transfer may be defined as a reaction in which one of the constituent groups of a condensed molecule, the donor, is transferred to new linkage with an acceptor. No change in the number of anhydro-bonds occurs during such reactions and, provided that the original and the final anhydro-bonds are of the same type, such reactions are accompanied by relatively small changes in the free energy; this is not the case, however, when very different types of anhydro-linkage are in question, as for example in the transfer of a phosphoryl group from pyrophosphate to ester linkage.

While it is often stated that the bond energy of the donor is used in the formation of the derived condensed product, a more concrete view of the chemical transformations is useful even though the underlying mechanism of the group transfer reaction, and the part played by the enzyme in the transformation (for example, whether or not the enzyme acts as intermediate acceptor and then donor of the group which is transferred), remain conjectural. The end result of a group transfer reaction, however, is the liberation of one of the groups from linkage in the donor and its replacement by the acceptor group. An essential feature of this change is that the liberated group will have gained one or other of the " elements " of water (H or OH), and that the acceptor will have lost the same " element " on entering anhydro-linkage. Participation of the new condensed molecule in a second transfer reaction can now lead to the liberation of the second group of the original donor; in this case the liberated group will have gained the other " element " of water, and the acceptor will have lost the same. The net result of the two reactions is the condensation of the two molecules which have acted as acceptors, with concomitant breakdown of the donor molecule. This breakdown is not due to hydrolysis in the ordinary sense, although the products are the same; the lysis will have resulted from participation not of water as such, but of the two elements of water derived separately from the two acceptors at the two stages of their condensation.

The group transfer reaction is thus seen to be a mechanism whereby one species of condensed molecule can be brought to function as an agent for the secondary production of other

species of condensed molecule. The role of ATP as a condensing agent in this sense is, of course, quite unique in the variety of such reactions in which it participates, in the close relationship of its generation with the primary condensation reactions, and, in many cases, in the relatively large decrease in free energy which accompanies the transfer of its terminal phosphoryl group to various acceptors. It is desirable therefore to consider almost as a separate category the condensations which are dependent upon concomitant breakdown of ATP, i.e. ATP-dependent condensations.

2. Synthesis of Peptide Bonds in ATP-Dependent Condensations

This avenue of investigation was stimulated by the suggestion of Lipmann (1941) that the condensation of carboxyl and amino groups of the amino-acids to form peptides might involve the intermediate formation of amino-acylphosphates generated by transfer from ATP.

Enzymic reactions were discovered in which the synthesis of certain peptide-like bonds occurs with concomitant breakdown of ATP. These are the acetylation of aromatic amines by acetate and ATP (Lipmann, 1945), and the synthesis of hippuric acid from benzoic acid, glycine, and ATP (Cohen & McGilvery, 1947). Perhaps of more direct relationship to the problem of protein metabolism was the discovery of the enzymic synthesis of glutamine from glutamic acid, ammonia, and ATP (Speck, 1947, 1949; Elliott, 1948).

More recently, the ATP-dependent synthesis of the tripeptide glutathione (γ-glutamylcysteinylglycine), catalysed by an enzyme system from liver tissue, has been discovered and elucidated in a series of investigations by Bloch and his colleagues (cf. recent review by Bloch, Snoke & Yanari, 1952). Evidence has been obtained that the synthesis takes place in two steps catalysed by different enzymes, but each being ATP-dependent. The dipeptide γ-glutamylcysteine is first synthesized, and this is condensed with glycine to yield glutathione in the second reaction.

It is to be noted that only in this last system has the synthesis of true peptide bonds been observed, and also that only the linkage between cysteine and glycine is of the type which occurs commonly, and perhaps exclusively, in proteins, i.e. α-carboxyl linked to α-amino groups.

While these investigations, culminating in the synthesis of glutathione, establish that certain peptide and amide linkages can be synthesized enzymically in ATP-dependent condensations, the range of amino-acids involved in the systems discovered to date is limited to glutamic acid, cysteine, and glycine.

3. Transpeptidation and Transamidation Reactions

A second field of investigation has been the search for transpeptidation and transamidation reactions, in which either the carboxyl or the amine moiety of a peptide or amide might be transferred to new linkage with the appropriate group of an acceptor amino-acid or peptide. Given mechanisms for the synthesis of particular peptides or amides in ATP-dependent condensations (e.g., glutamine, glutathione), the secondary derivation from these of varied peptides and amino-acid sequences might be envisaged as resulting from appropriate sequences of transfer reactions.

Such was the speculative conception which led our group at Cambridge (F. J. R. Hird, F. A. Isherwood, and the writer) to undertake a search for enzymes catalysing such reactions in extracts from plant and animal sources. At about the same time, J. S. Fruton at Yale University had embarked upon allied investigations which were prompted by very similar considerations (cf. Fruton, 1950; Hanes, Hird & Isherwood, 1950).

In view of the unique position of glutathione as a small peptide of widespread distribution in tissues, coupled with the fact that it contained the γ-glutamyl group, the condensation of which in glutamine had been established, and finally the suggestion that glutathione may act in some way as an intermediary between amino-acids and proteins (Waelsch & Rittenberg, 1941), our attention was turned at the outset to the possibility that this peptide might participate in transpeptidation reactions. Clear indications that this was the case were soon obtained in experiments with kidney extracts carried out by Dr Hird, and these led to the recognition of the enzyme to be referred to as γ-glutamyl transpeptidase. A second enzyme from a plant source was found to catalyse the transfer of the glycyl group from glycyl peptides to various amino-acid acceptors. A considerable variety of transpeptidation reactions catalysed by these two enzymes have been described (Hanes, Hird & Isherwood, 1950, 1952b) and observations upon the γ-glutamyl transpeptidase have been reported by other workers (cf. Waelsch, 1952a, 1952b; Kinoshita & Ball, 1953).

In the meantime Fruton and various colleagues (Fruton, 1950; Johnston, Mycek & Fruton, 1950a, 1950b; Fruton, Johnston & Fried, 1951; Fruton, 1952) have described a considerable variety of transpeptidation and transamidation reactions which they have found to be catalysed by various of the familiar proteolytic enzymes.

Since detailed reviews of transpeptidation and transamidation reactions have appeared recently (Hanes, Connell & Dixon, 1952a; Fruton, 1952; Waelsch, 1952a, 1952b), we need concern ourselves only with certain general features of these reactions.

a. Transfers of α-Amino-Acyl Groups

The reactions described by Fruton and his colleagues consist in all cases of transfers of α-amino-acyl groups from either amide or peptide linkage. Reactions of this type were found to be catalysed by the following enzymes: papain, ficin, trypsin, chymotrypsin, and two cathepsins from beef-spleen tissue.

All these enzymes were shown to catalyse transfer of appropriate groups (i.e. consistent with the requirements imposed by the substrate specificity of the proteolytic enzymes) from amide linkage to such acceptors as hydroxylamine, amino-acid amides, and (in certain cases) amides of di- and oligo-peptides. The groups transferred were most frequently *N*-substituted amino-acyl groups (e.g., benzoylglycyl; carbobenzoxy-L-α-glutamyl), but instances of the transfer of dipeptide and *N*-substituted dipeptide groups are also reported.

It appears from a survey of the reactions described that the specificity of each enzyme in regard to group transfer and hydrolysis is the same.

In a number of the systems reported upon, the formation of polypeptide chains was observed, these being the result of repeated transfers leading to the elongation of the resulting

condensed product. An example of this type (Jones *et al.* 1952) occurred as a result of the action of cathepsin C on glycyl-phenylalanylamide, during which a gelatinous precipitate was formed which was identified as a mixture of an octapeptide and its amide with the probable structure (glycyl-phenylalanyl-)$_4$-amide.

Other examples of the formation of longer-chain polymers have been reported briefly by Fruton (1952) and in some cases under suitable conditions such products have been obtained in yields amounting to 70% of the maximum theoretically possible.

b. Transfers of γ-Glutamyl Groups

The γ-glutamyl transpeptidase, first observed in extracts of mammalian kidney and pancreas, was found to catalyse transfers of the γ-glutamyl group from glutathione to a number of amino-acids as acceptors, with the formation of the corresponding γ-glutamyl peptides. Later it was found that glutathione is not unique as a donor of γ-glutamyl groups in this system, but various synthetic γ-glutamyl peptides served equally well. The effect of the transfer reaction is to bring about the distribution of the γ-glutamyl group between different potential acceptors present in the system (Hanes *et al.* 1950, 1952b). Waelsch (1952b) reported the interesting reaction in which the γ-glutamyl group was transferred from γ-glutamylglycine, as donor, to cysteinylglycine with the resultant formation of glutathione. Such a mechanism of synthesis of glutathione had been envisaged earlier in connexion with two hypotheses. The first was the possible synthesis of glutathione by the corresponding transfer from glutamine (Hanes *et al.* 1950), a reaction which has not been realized because the enzyme has so far failed to act upon glutamine as donor of γ-glutamyl groups. The second hypothesis hinges on the fact that the cysteinylglycine bond in glutathione is not hydrolysed by the widespread peptidase which induces rapid splitting of free cysteinylglycine (cf. Binkley & Nakamura, 1948). Accordingly, transfer of γ-glutamyl group to the dipeptide would remove it as glutathione from the sphere of action of the peptidase; therefore synthesis of glutathione might be expected to occur in a system containing cysteine, glycine, a donor of γ-glutamyl groups (e.g., γ-glutamylglutamic acid) and the two enzymes. The mechanism of this postulated reaction would be direct condensation (reversal of hydrolysis) of cysteine and glycine, promoted by the removal of the product from the action of the peptidase by its conversion into glutathione.

Reference should now be made to a group of enzymes, widespread in distribution, which have been termed glutamotransferases (cf. Waelsch, 1952a, 1952b) and which catalyse the transfer of the γ-glutamyl group from glutamine to hydroxylamine or to isotopic ammonia, but not, so far as is known, to any amino-acid or other acceptor. These enzymes, which were first discovered by Waelsch in micro-organisms and later in mammalian tissues, have been found also in higher plants by Stumpf & Loomis (1950). The enzymes from the different sources are strikingly different as regards their co-factor requirements, but they show in common the curiously restricted range in the reactions they are able to catalyse. It seems probable that they may be found to possess other catalytic properties, which may make it possible to envisage their biological significance.

4. Discussion

It will be evident from the above survey that important progress has been made in recent years in the discovery of different types of enzymic reactions which lead to the formation of peptides. The question naturally arises whether these reactions form parts of the mechanism of protein synthesis in the cell. The answer must be that in no case has this been clearly established. But the present body of experimental facts is still very limited in relation to the magnitude of the problem as a whole, and until the main lines of the synthetic pathway are established there can be no proof that this or that particular type of reaction forms an essential part of the whole process. It can be said, however, that the progress which is being made is encouraging. Indeed it is tempting to formulate speculative hypotheses incorporating suggestions as to how the reactions now recognized may form parts of a unified mechanism; such formulations at the present juncture would almost certainly include suggestions as to how glutamine and glutathione, which are formed in ATP-dependent condensations, might take part in transpeptidation reactions leading to the formation of peptide-linked chain molecules.

While the value of such speculations is clear, it may be more profitable to take note of some of the obvious gaps in existing knowledge. Perhaps the most important question is one arising from the discovery of systems synthesizing glutamine and glutathione: it is whether analogous enzymic mechanisms exist whereby some wider range of amino-acids can be condensed in ATP-dependent reactions. It would seem unlikely that the two systems so far discovered represent the only examples of their type, and the search for other such reactions should have high priority.

Consideration of the available information about transpeptidation reactions leads to other questions on which further knowledge may be of assistance in defining the part which these reactions may play in the elaboration and transformation of peptide structure.

It should be noted, for example, that—with the exception of the group of enzymes that have been termed glutamotransferases—all the remaining enzymes which have been found to catalyse transpeptidation and transamidation reactions are able also to catalyse the hydrolysis of the linkages upon which they act. This raises the question whether or not there may exist also specialized transferring enzymes which catalyse transpeptidations without hydrolysis; such specialized transferring enzymes are known in the fields of transphosphorylation and transglycosidation reactions. Whether or not such enzymes exist, there are interesting indications from several of the transpeptidation systems which have been studied that the ratio of hydrolysis to group transfer may be considerably influenced by various factors. Further information on these effects, especially of quantitative character, would be most valuable.

In this connexion it is clear that quantitative data permitting the clarification of kinetic aspects of transpeptidation systems would be of great interest. The scarcity of such information at present has been due to the analytical difficulties inherent in the separation and determination of closely related peptides in the presence of their constituent amino-acids. There are indications, however, that the satisfactory solution of these problems will not be long delayed and much important information should then become available.

A transpeptidation reaction might theoretically involve the transfer of either the carboxyl or the amino moiety of the donor to new linkage with the appropriate group of the acceptor. It should be noted that all the examples so far observed are of the carboxyl transfer type. The question whether or not reactions of the alternative (amino transfer) type also occur is likely to become important in delineating the role of transpeptidation reactions as elements in synthetic pathways.

ACKNOWLEDGEMENT

The writer carried out his work in this field under the aegis of the Agricultural Research Council whilst he was in England, and he is glad here to record his appreciation of the unfailing support he received from the Council. The continuation of the work in related fields at the University of Toronto is being supported generously by the J. P. Bickell Foundation.

REFERENCES

Binkley, F. & Nakamura, K. (1948) *J. biol. Chem.* **173,** 411

Bloch, K., Snoke, J. E. & Yanari, S. (1952) In: McElroy, W. D. & Glass, H. B., ed. *Phosphorus metabolism: a symposium on the role of phosphorus in the metabolism of plants and animals,* vol. 2, p. 82. Johns Hopkins Press, Baltimore

Cohen, P. P. & McGilvery, R. W. (1947) *J. biol. Chem.* **169,** 119

Elliott, W. H. (1948) *Nature, Lond.* **161,** 128

Fruton, J. S. (1950) *Yale J. Biol. Med.* **22,** 263

Fruton, J. S. (1952) *II^e Congrès International de Biochimie, Paris, 21–27 Juillet 1952.* In : *Symposium sur la biogénèse des protéines,* p. 5. Société d'Edition d'Enseignement Supérieur, Paris

Fruton, J. S., Johnston, R. B. & Fried, M. (1951) *J. biol. Chem.* **190,** 39

Hanes, C. S., Connell, G. E. & Dixon, G. H. (1952a) In: McElroy, W. D. & Glass, H. B., ed. *Phosphorus metabolism: a symposium on the role of phosphorus in the metabolism of plants and animals,* vol. 2, p. 95. Johns Hopkins Press, Baltimore

Hanes, C. S., Hird, F. J. R. & Isherwood, F. A. (1952b) *Biochem. J.* **51,** 25

Hanes, C. S., Hird, F. J. R. & Isherwood, F. A. (1950) *Nature, Lond.* **164,** 1107

Harden, A. & Young, W. J. (1905) *Proc. chem. Soc., Lond.* **21,** 189

Johnston, R. B., Mycek, M. J. & Fruton, J. S. (1950a) *J. biol. Chem.* **185,** 629

Johnston, R. B., Mycek, M. J. & Fruton, J. S. (1950b) *J. biol. Chem.* **187,** 205

Jones, M. E., Hearn, W. R., Fried, M. & Fruton, J. S. (1952) *J. biol. Chem.* **195,** 645

Kinoshita, J. H. & Ball, E. G. (1953) *J. biol. Chem.* **200,** 609

Lipmann, F. (1941) *Advanc. Enzymol.* **1,** 99

Lipmann, F. (1945) *J. biol. Chem.* **160,** 173

Speck, J. F. (1947) *J. biol. Chem.* **168,** 403

Speck, J. F. (1949) *J. biol. Chem.* **179,** 1405

Stumpf, P. K. & Loomis, W. D. (1950) *Arch. Biochem.* **25,** 451

Waelsch, H. (1952a) *II^e Congrès International de Biochimie, Paris, 21–27 Juillet 1952.* In : *Symposium sur la biogénèse des protéines,* p. 26. Société d'Edition d'Enseignement Supérieur, Paris

Waelsch, H. (1952b) In : McElroy, W. D. & Glass, H. B., ed. *Phosphorus metabolism: a symposium on the role of phosphorus in the metabolism of plants and animals,* vol. 2, p. 109. Johns Hopkins Press, Baltimore

Waelsch, H. & Rittenberg, D. (1941) *J. biol. Chem.* **139,** 761

* * *

AMINO-ACID DECARBOXYLASES

E. F. GALE Sc.D. F.R.S.

Director, Medical Research Council Unit for Chemical Microbiology

University Reader in Chemical Microbiology University of Cambridge

1 Bacterial amino-acid decarboxylases
- *a* Growth conditions necessary for production of active cells
- *b* Specific decarboxylases identified
- *c* Use of enzyme preparations for assay purposes
- *d* Properties of the enzymes: nature of prosthetic group

2 Amino-acid decarboxylases of other cells
3 Pathogenic significance
References

Amines have long been known as products of bacterial putrefaction and at the beginning of the present century a number of workers showed that amino-acids were the natural precursors of such amines. The early work was carried out under conditions of "controlled putrefaction" in which media containing salts and certain known amino-acids were inoculated with putrefying material, allowed to incubate for weeks or months, and the products of putrefaction then isolated and identified. In this way the microbial formation of putrescine, histamine, cadaverine, tyramine and agmatine from the corresponding amino-acids was demonstrated. It was also found that aspartic acid or glutamic acid, under similar conditions, gave rise to β-alanine or γ-aminobutyric acid respectively, and the suggestion was clear that mixed cultures of micro-organisms are able to remove the carboxyl group from certain amino-acids with the production of the corresponding amine or, in the case of the dicarboxylic acids, the corresponding monocarboxylic amino-acid. Pure strains of intestinal bacteria were shown to produce histamine from histidine (Mellanby & Twort, 1912), while Berthelot & Bertrand (1912) isolated an organism capable of forming the corresponding amines from histidine, tyrosine and tryptophan.

Later workers attempted to define more closely the conditions in which bacteria would decarboxylate amino-acids. Especial reference was paid to histamine formation and the general conclusion arrived at by Koessler & Hanke (1919) was that "histamine is never formed except in the presence of an easily available source of carbon such as glycerol or glucose". Eggerth (1939) found that histamine formation by a variety of bacteria depended not only on the presence of such available sources of carbon but on the pH of the medium in which the cells were growing, and it became apparent that the carbon sources previously used were acting more as a source of fermentation acid than as a source of carbon as such. Gale (1940) found that washed suspensions of bacterial cells (*Bacterium coli, Streptococcus faecalis, Clostridium* spp., etc.) could directly and rapidly decarboxylate certain amino-acids if the cells had first been cultured under defined conditions, important amongst which were (i) pH well on the acid side of neutrality and (ii) the presence in the medium of the particular amino-acid or acids concerned. Later Gale & Epps (1944) found that the decarboxylation reactions were carried out by specific enzymes which could be extracted from the cells. A number of such enzymes were isolated and partially purified; each was found to decarboxylate a specific amino-acid with the quantitative formation of the corresponding amine, and these enzymes were called the "amino-acid decarboxylases". Other studies have shown the occurrence of similar activities in plants (Okunuki, 1937; Schales, Mims & Schales, 1946) and in mammalian tissues (Werle, 1936; Holtz & Heise, 1937; Blaschko, 1945). Detailed reviews of the subject have been written by Blaschko (1945), Gale (1946) and Schales (1951).

1. Bacterial Amino-Acid Decarboxylases

a. Growth Conditions Necessary for Production of Active Cells

The ability to form amino-acid decarboxylases is not possessed by all bacteria. Table I outlines some of the common organisms that have been investigated and the decarboxylases that they are able to produce when grown under the right conditions. Wide strain differences are

TABLE I. DISTRIBUTION OF AMINO-ACID DECARBOXYLASES]

Source	Amino-acid substrate										
	Histidine	Arginine	Lysine	Ornithine	Tyrosine	Glutamic acid	Aspartic acid	Cysteic acid	Phenylalanine	Tryptophan	Hydroxy-phenylserine
Bact. coli	+	+	+	+	+	+	+		−	−	
Str. faecalis	−	−	−	−	+	−	−		+	−	
Cl. welchii	+	−	−	−	−	+	+		−	−	
Cl. fallax	+	−	−	−	−	−	−		−	−	
Proteus vulgaris	−	−	−	+	−	+	−		−	−	
Rhizobium trifolium	−	−	−	−	−	−	+		−	−	
Lactobacillus spp.	+	+	+	+	−	+	−		−	−	
Squash seed						+					
Carrot						+					
Cucumber						+					
Guinea-pig liver	+				+			+	+	+	+
Mouse brain						+					

encountered within species: thus strains of *Bact. coli* may form any number from 0 to 6 of the decarboxylases listed. The first essential is therefore that the organism chosen shall be able to form the enzyme; then it must be grown under the right conditions to evoke such formation. Many of the decarboxylases that have been investigated prove to be adaptive enzymes[1]; consequently they are produced only if the organism is grown in the presence of the specific substrate. Secondly the adaptation responds to the presence in the medium of the amino-acid in the cationic state only; this requires that the medium shall be acid. In general the greatest production of decarboxylases occurs when growth takes place at a pH value near the limit of acid tolerance for the organism concerned, but the amount of growth under these conditions is small. Heavy growth of active organisms can, however, be obtained if the medium contains fermentable carbohydrate which then acts as a source both of carbon and energy for growth, and of acid for stimulating decarboxylase formation. Cells do not produce decarboxylases readily during the early stages of growth but contain the greatest amount of enzyme if they are harvested towards the end of the period of active growth; the activity may decline rapidly once growth has ceased. The prosthetic group of many of these enzymes is pyridoxal phosphate and, if the organism cannot readily synthesize this substance, it (or its precursors) must be supplied in the medium. A curious finding is that the amount of pyridoxine which must be supplied in the medium in some cases (e.g., the formation of ornithine decarboxylase by *Lactobacillus* spp. (Rodwell, 1953)) is many times greater than the amount required to give optimum growth of the organism.

When all these conditions are fulfilled, decarboxylation of the amino-acid substrate in the medium occurs, some neutralization of the acidity of the medium takes place, and the harvested cells are found to be highly active as preparations of the amino-acid decarboxylase concerned.

The properties of diaminopimelic acid decarboxylase, recently described in *Bact. coli* by Dewey & Work (1952), differ in some respects from those described above. Thus this enzyme appears to be constitutive (as opposed to adaptive) and is optimally active near neutrality. Its formation is greater during growth in acid growth conditions but does not appear to be seriously affected by the temperature at which growth takes place, whereas the formation of other decarboxylases in *Bact. coli* is often markedly more efficient at temperatures of 20°–30° C. than at higher values.

b. Specific Decarboxylases Identified

The distribution of decarboxylating activities amongst bacterial species indicates that there is no constant association of activity towards any one amino-acid with that towards any other. Extraction of the enzymes from the cells, followed by partial purification, has shown that each enzyme is specific for one amino-acid substrate. Enzymes have been isolated which are specific for the decarboxylation of lysine, arginine, ornithine, histidine, glutamic acid, aspartic acid, tyrosine, diaminopimelic acid and, possibly, valine. In each case the enzyme attacks only the natural optical isomer, and any masking of the polar groups within the substrate renders the substance resistant to attack by the enzyme. For example, in the case of lysine decarboxylase, the enzyme cannot attack lysine in which a substituent has been introduced into either of the amino groups or the carboxyl group. Consequently the enzymes will not attack amino-acid residues occurring as peptides or proteins. Some degree of modification of the structure of the substrate can, however, take place without inactivation of the decarboxylation reaction. Thus preparations of lysine decarboxylase will attack hydroxylysine, glutamic acid decarboxylase will attack β-hydroxyglutamic acid, while tyrosine decarboxylase preparations will attack phenylalanine, *m*-tyrosine, 3 : 4-dihydroxyphenylalanine and 2 : 5-dihydroxyphenylalanine. It is not yet known whether these activities are due to the same enzyme in each case as tests have not been carried out with pure enzyme preparations. Consequently the enzyme which attacks glutamic acid may not be the same as that which decarboxylates β-hydroxyglutamic acid, whilst the decarboxylation of phenylalanine derivatives may be due to the action of a number of closely associated but specific enzymes. In all cases except that of aspartic acid decarboxylation, the action of the enzyme is to remove the —COOH group next to the α-amino group. In the case of aspartic acid, it appears that two decarboxylases may exist: decarboxylation in the usual way would yield β-alanine as the product and this activity has been demonstrated in suspensions of *Rhizobium trifolium* (Virtanen & Laine, 1937; Billen & Lichstein, 1949), but Meister, Sober & Tice (1951) have found that suspensions of *Clostridium welchii*, in the presence of catalytic amounts of an α-keto acid, accomplish a decarboxylation of aspartic acid to yield α-alanine. Decarboxylation of diaminopimelic acid yields lysine (Dewey & Work, 1952) and it is probable that diaminopimelic acid is the biosynthetic precursor of lysine in organisms possessing this enzyme (Davis, 1952).

c. Use of Enzyme Preparations for Assay Purposes

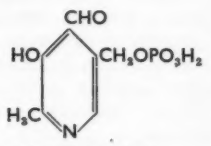
Since the action of these enzymes is specific and quantitative, they can be used manometrically for the rapid and accurate determination of their substrates (Gale, 1945a, 1945b). They have been used extensively for amino-acid estimations in protein hydrolysates and in studies of amino-acid metabolism in micro-organisms.

d. Properties of the Enzymes : Nature of Prosthetic Group

Although none of these enzymes has yet been obtained in a completely pure condition, partially purified preparations of most of them have been made and their properties studied. With the exception of the diaminopimelic acid decarboxylase, the enzymes have unusually acid pH optima. Thus histidine decarboxylase is optimally active at pH 4.0–4.5 and glutamic acid decarboxylase at pH 4.5. This accords with the acid growth conditions necessary for their formation. The decarboxylation proceeds to completion in all cases examined and no convincing demonstration of reversibility has been made. It has not been possible to show transfer of isotopic carbon into lysine during decarboxylation of this amino-acid in an atmosphere of labelled carbon dioxide (Werkman, private communication). It seems improbable therefore that these enzymes can be concerned in the synthesis of amino-acids from the corresponding amines. In this connexion it is interesting that a strain of *Str. faecalis*, which is nutritionally exacting towards phenylalanine and which contains phenylalanine decarboxylase, can grow in the presence of phenylethylamine as long as carbon dioxide is supplied (Lyman *et al.* 1947).

[1] See Hinshelwood, p. 150 of this Bulletin.

Studies on the formation of tyrosine decarboxylase by *Str. faecalis* showed that active cells were obtained only if growth took place in the presence of excess pyridoxine (Bellamy & Gunsalus, 1944). Gale and co-workers (see Gale, 1946) showed that it is possible to inactivate preparations of either lysine or tyrosine decarboxylase by treatment with ammoniacal ammonium sulphate solution, and that such inactivated preparations (apoenzyme) could be re-activated by the addition of a thermostable coenzyme present in all living tissues examined. Preparations of tyrosine apodecarboxylase or of pyridoxine-deficient cells could be activated by derivatives of pyridoxine. To obtain active codecarboxylase from pyridoxine, it was necessary first to oxidize it to pyridoxal and then to phosphorylate the pyridoxal. Some doubt has existed concerning the correct formula of "pyridoxal phosphate" but Baddiley & Mathias (1952) have recently accomplished a synthesis which is put forward as proof of the following structure:



pyridoxal phosphate

Studies of the bacterial amino-acid decarboxylases first demonstrated the existence of pyridoxal phosphate, the functional form of vitamin B_6, and preparations of the apodecarboxylases provide means for the assay of this coenzyme. Pyridoxal phosphate is now known to have many functions other than its action as codecarboxylase, and forms the prosthetic group of many enzymes involved in the transamination, synthesis and breakdown of amino-acids. With the exception of histidine decarboxylase, all the bacterial amino-acid decarboxylases so far studied require pyridoxal phosphate as prosthetic group. It has not been possible to demonstrate activation of histidine decarboxylase by pyridoxal phosphate, and Rodwell (1953) has shown that lactobacilli, grown in media so deficient in pyridoxine that other decarboxylases are completely inactive, possess greater histidine decarboxylase activity than when grown in pyridoxine-rich media. Werle (1947–48) has found that the histidine decarboxylase of mammalian tissues is inhibited by pyridoxal phosphate.

2. Amino-Acid Decarboxylases of Other Cells

In general, bacteria form the most active sources of amino-acid decarboxylases and, on a dry-weight basis, may have activities 100–1,000 times those of plant and animal tissues. However, a potent glutamic acid decarboxylase is found in a variety of plant tissues (Schales *et al.* 1946) and a highly active preparation for analytical purposes can be made from squash seeds (Schales & Schales, 1946). The pH optimum for the squash glutamic acid decarboxylase is 6.0 compared with 4.5 for the bacterial enzyme, while the similar enzyme found in brain is optimally active at 6.5 (Roberts & Frankel, 1951).

Table I shows that enzymes which attack substrates such as tryptophan, cysteic acid and hydroxyphenylserine are found in liver cells. Decarboxylation of these substances has not been demonstrated in bacterial preparations although the activities found in mammalian tissues would probably be too small to have been detected by the methods used for the study of bacterial enzymes. Tryptamine was, however, present amongst the products of bacterial putrefaction found in experiments carried out at the beginning of the century. Blaschko (1942) demonstrated the occurrence of cysteic acid decarboxylase in liver. The product of this enzyme is taurine, and, since Bernheim & Bernheim (1939) had previously shown the formation of cysteic acid from cysteine in the liver, it is now possible to trace the course of the production of taurine in that tissue.

One of the more active decarboxylases of mammalian tissues is the dihydroxyphenylalanine (DOPA) decarboxylase studied in detail by Blaschko & Sloane Stanley (1948) at Oxford. Some interesting differences in the specificities of the bacterial and guinea-pig preparations were observed in that the bacterial preparation attacked 4-hydroxyphenylalanine, whereas the kidney preparation did not, while the reverse situation was found for 2-hydroxyphenylalanine and 2:5-dihydroxyphenylalanine.

3. Pathogenic Significance

Histamine is stored by many tissues and the release of histamine during tissue damage arises from such stored material rather than from decarboxylation of histidine. The activity of the histidine decarboxylase of mammalian tissues is so small that its significance in histamine release is difficult to assess (see Blaschko, 1945).[1] Histamine formation by bacteria in vivo is of possible significance. When *Cl. welchii* infection gives rise to gas gangrene, conditions of amino-acid release and acid formation occur in wounds which might be suitable for histamine production by the histidine decarboxylase of the infecting organism. Brown, MacIntosh & Bruce White (1941) investigated the histamine content of muscles of cats infected with *Cl. welchii* and found that it rose to a level 30–250% higher than that of control uninfected muscles. The amounts found, however, were so small that the authors concluded that no significant absorption into the blood stream would result.

Åkerblom (1934) investigated the possibility that histamine intoxication was responsible for *Futterrehe* in horses. He found that the symptoms could be produced by injection of histamine, or by feeding diets high in histidine. He demonstrated the presence of histamine in stomach and intestinal contents and isolated coliform organisms therefrom which were capable of decarboxylating histidine. In the course of the studies at Cambridge (see Gale, 1946), some hundreds of coliform organisms were investigated and, although histidine decarboxylase was frequently encountered in these organisms, the activity was always small. However, Rodwell (1953) has isolated a new species of *Lactobacillus* from the stomach contents of horses; these bacteria have histidine decarboxylase activities 30–50 times greater than that of the average coliform organism, and ten times higher than that of the most active *Cl. welchii* strains isolated. These organisms can produce 1 mg. of histamine per millilitre of culture medium and it seems feasible that they might produce sufficient histamine in vivo to have pathogenic actions.

The other biologically produced amines have comparatively little pharmacological activity, and the amine oxidase

[1] See also Blaschko, p. 146 of this Bulletin.

Continued at foot of page 138

CHOLINESTERASES

R. H. S. THOMPSON M.A. D.M.

Department of Chemical Pathology
Guy's Hospital Medical School, London

1 Types of cholinesterase
2 Inhibitors of cholinesterase
3 Limitations of present classifications of cholinesterases
4 Distribution of cholinesterase types in different tissues
5 Effects of inhibition of cholinesterase
6 Physiological functions of cholinesterases
References

Interest in the group of enzymes now known as the cholinesterases may be said to have started in 1914 when Sir Henry Dale first suggested that the transient action of acetylcholine in vivo was probably due to its destruction by an esterase in the blood. Twelve years later, in 1926, Loewi & Navratil succeeded in demonstrating the presence of such an enzyme in tissue extracts, and in 1932 the term "choline-esterase" was introduced by Stedman, Stedman & Easson to describe enzymes that hydrolyse acetylcholine. It soon became apparent, however, that the cholinesterases from different sources, and even from different tissues in the same animal, show striking differences in properties. Despite much work, the specificity of these and other esterases remains an outstanding problem in this field, with which is linked the broader question of their physiological functions.

1. Types of Cholinesterase

Alles & Hawes (1940) demonstrated that the cholinesterase of human erythrocytes is much more active in hydrolysing acetylcholine at low substrate concentrations than is the enzyme in human plasma. They also described differences in substrate specificity between the cholinesterases in these two tissues, acetyl-α-methylcholine being hydrolysed by both enzymes, but acetyl-β-methylcholine only by the erythrocyte enzyme. Further differences in the specificity and kinetics of cholinesterases from different sources were described by Glick (1941), Richter & Croft (1942) and Zeller & Bissegger (1943). Mendel and his colleagues in Toronto next proposed a classification of the cholinesterases into two main types, which they termed "true" and "pseudo" cholinesterase (Mendel & Rudney, 1943a). The "true" cholinesterase, present in erythrocytes and in the nervous system, was thought (i) to hydrolyse only choline esters, (ii) to be partially inhibited by high concentrations of acetylcholine, and (iii) to hydrolyse acetyl-β-methylcholine but not benzoylcholine; "pseudo-cholinesterase", present in the serum and the pancreas, was thought (i) to hydrolyse both choline and non-choline esters, (ii) to show maximal activity at high acetylcholine concentrations, and (iii) to hydrolyse benzoylcholine but not acetyl-β-methylcholine.

2. Inhibitors of Cholinesterase

Mendel's proposal of these two main types of cholinesterase received further support when the effect of inhibitors on cholinesterase activity was studied. Use had been made for many years of the inhibition of cholinesterase activity by eserine (Englehart & Loewi, 1930) and by prostigmine (Easson & Stedman, 1937), and, in general, these are highly effective inhibitors for most of the mammalian cholinesterases, irrespective of their type according to Mendel's classification. A large and rapidly growing series of compounds, however, is

---

AMINO-ACID DECARBOXYLASES
Continued from page 137

system of mammalian tissues can effectively destroy such quantities of these amines as are likely to occur naturally. However, newly born infants are deficient in both tyramine oxidase (Epps, 1945) and histaminase (Zeller *et al.* 1939; Zeller, Stern & Wenk, 1940) and it is possible that amine intoxication might assume importance in infections of the new-born.

REFERENCES

Åkerblom, E. (1934) *Skand. Arch. Physiol.* **68**, Suppl.
Baddiley, J. & Mathias, A. P. (1952) *J. chem. Soc.*, p. 2583
Bellamy, W. D. & Gunsalus, I. C. (1944) *J. Bact.* **48**, 191
Bernheim, F. & Bernheim, M. L. C. (1939) *J. biol. Chem.* **127**, 695
Berthelot, A. & Bertrand, D. M. (1912) *C.R. Acad. Sci., Paris*, **154**, 1643; 1826
Billen, D. & Lichstein, H. C. (1949) *J. Bact.* **58**, 215
Blaschko, H. (1942) *Biochem. J.* **36**, 571
Blaschko, H. (1945) *Advanc. Enzymol.* **5**, 67
Blaschko, H. & Stanley, G. H. Sloane (1948) *Biochem. J.* **42**, *Proc.* iii
Brown, G. L., MacIntosh, F. C. & White, P. Bruce (1941) *Biochem. J.* **35**, 79
Davis, B. D. (1952) *Nature, Lond.* **169**, 534
Dewey, D. L. & Work, E. (1952) *Nature, Lond.* **169**, 533
Eggerth, A. H. (1939) *J. Bact.* **37**, 230
Epps, H. M. R. (1945) *Biochem. J.* **39**, 37
Gale, E. F. (1940) *Biochem. J.* **34**, 392; 846; 853
Gale, E. F. (1945a) *Biochem. J.* **39**, 46
Gale, E. F. (1945b) *Nature, Lond.* **157**, 265
Gale, E. F. (1946) *Advanc. Enzymol.* **6**, 1
Gale, E. F. & Epps, H. M. R. (1944) *Biochem. J.* **38**, 232
Holtz, P. & Heise, R. (1937) *Arch. exp. Path. Pharmak.* **186**, 377
Koessler, K. K. & Hanke, M. T. (1919) *J. biol. Chem.* **39**, 539
Lyman, C. M., Moseley, O., Wood, S., Butler, B. & Hale, F. (1947) *J. biol. Chem.* **167**, 177
Meister, A., Sober, H. A. & Tice, S. V. (1951) *J. biol. Chem.* **189**, 577
Mellanby, E. & Twort, F. W. (1912) *J. Physiol.* **45**, 53
Okunuki, C. (1937) *Botan. Mag., Tokyo*, **51**, 270
Roberts, E. & Frankel, S. (1951) *J. biol. Chem.* **190**, 505
Rodwell, A. W. (1953) *J. gen. Microbiol.* **8**, 224; 233; 238
Schales, O. (1951) In: Sumner, J. B. & Myrbäck, K. D. R., ed. *The enzymes: chemistry and mechanism of action*, vol. 2, pt 1, p. 216. Academic Press, New York
Schales, O., Mims, V. & Schales, S. S. (1946) *Arch. Biochem.* **10**, 455
Schales, O. & Schales, S. S. (1946) *Arch. Biochem.* **11**, 445
Virtanen, A. I. & Laine, T. (1937) *Enzymologia*, **3**, 266
Werle, E. (1936) *Biochem. Z.* **288**, 292
Werle, E. (1947–48) *Z. Vitam.- Horm.- u. Fermentforsch.* **1**, 504
Zeller, E. A., Birkhauser, H., Mislin, H. & Wenk, M. (1939) *Helv. chim. Acta*, **22**, 1381
Zeller, E. A., Stern, R. & Wenk, M. (1940) *Helv. chim. Acta*, **23**, 3

now known which selectively inhibit either true cholinesterase or the pseudo-cholinesterase. Among the earlier compounds to be studied, caffeine (Zeller & Bissegger, 1943), bis-β-chloroethyl-*N*-methylamine, or "nitrogen mustard" (Thompson, 1947; Adams & Thompson, 1948), and the synthetic prostigmine analogue Nu 1250 (*N*-*p*-chlorophenyl-*N*-methylcarbamate of *m*-hydroxyphenyl-trimethylammonium bromide) are all selective inhibitors of the true cholinesterase. Percaine (Zeller & Bissegger, 1943), tri-*o*-cresyl phosphate (Mendel & Rudney, 1944; Earl & Thompson, 1952), di-*iso*propyl fluorophosphonate, or DFP (Mazur & Bodansky, 1946; Hawkins & Mendel, 1947) and another prostigmine analogue Nu 683 (the dimethylcarbamate of 2-hydroxy-5-phenylbenzyl-trimethylammonium bromide) all selectively inhibit the pseudo-cholinesterase.

Although tri-*o*-cresyl phosphate was first shown to inhibit horse serum cholinesterase by Bloch in 1941, the introduction of DFP, whose pharmacological action was described by Adrian, Feldberg & Kilby (1947), together with the development of the alkylpolyphosphates and aryldialkylphosphates as insecticides, has led to the intensive study of the group of organo-phosphorus anticholinesterases.

Many of these compounds are highly effective in extremely low concentrations; for example, tetraethyl pyrophosphate (TEPP) is an effective inhibitor in concentrations as low as 10^{-10} M. They differ from eserine (Matthes, 1930) or prostigmine, in that they appear to inhibit cholinesterase irreversibly (Mackworth & Webb, 1948) by reacting stoicheiometrically with the active centres of the enzyme. It is thought that a phosphonium ion $[(RO)_2PO]^+$ is split off, and that this then combines irreversibly with part of the active centre. The cholinesterase itself catalyses the formation of the ion which ultimately destroys its activity; as Whittaker (1951) has put it, the enzyme "commits suicide". The kinetics of the inhibition have been studied in detail by Aldridge (1950), who believes that two reactions take place:

$$E + I \rightleftharpoons EI \rightarrow EI'$$

first, the formation of a dissociable enzyme-inhibitor complex, *EI*, which then undergoes a unimolecular transformation into an undissociable complex, *EI'*.

3. Limitations of Present Classifications of Cholinesterases

Although Mendel's division of these esterases into the two types, true and pseudo-cholinesterase, has been of great value in that, in many cases, it allowed a greater precision in the definition of the enzyme concerned in acetylcholine hydrolysis in a given tissue, more recent work has shown that there are certainly more than two types of cholinesterase, as judged by the criteria usually employed in defining the true and pseudo types. For example, it soon became clear that many so-called true cholinesterases (or "specific" according to the nomenclature introduced by Nachmansohn & Rothenberg, 1945) hydrolyse triacetin more rapidly than acetylcholine, and that this triacetin hydrolysis is highly sensitive to anticholinesterases and is present in preparations freed from aliesterase (Mazur & Bodansky, 1946; Holton, 1948; Blaschko & Holton, 1949). Cholinesterases that do not fulfil the original criteria regarding activity towards "selective" substrates have also been described. Thus, swine serum contains a cholinesterase which hydrolyses neither acetyl-β-methylcholine nor benzoylcholine (Levine & Suran, 1950), while chicken serum cholinesterase hydrolyses both of these so-called selective substrates at quite rapid rates (Earl & Thompson, 1952).

In view of these findings, Adams and Whittaker in 1948 commenced a systematic survey of the substrate specificity of partially purified true and pseudo-cholinesterase. They confirmed the earlier claims that true or "specific" cholinesterases are capable of hydrolysing many non-choline esters. Their study of the effect of variations in the number of carbon atoms in the acyl group of the esters yielded some interesting results; they found that, both for choline esters and for aliphatic esters, the true cholinesterase hydrolyses acetates more rapidly than propionates, and butyrates only very slowly indeed (Adams & Whittaker, 1949a); with pseudo-cholinesterase, on the other hand, butyrates are hydrolysed faster than propionates, and acetates more slowly still (Adams & Whittaker, 1949b). Changes in the configuration of the alkyl group of the ester molecule were also of interest, in that 3:3-dimethylbutyl acetate, the carbon analogue of acetylcholine, was hydrolysed very much more rapidly than the other esters studied, both by true and pseudo-cholinesterases. Thus, both types of enzyme are "unspecific" (Nachmansohn & Rothenberg, 1945) in that they both hydrolyse non-choline esters, although with both enzymes there is a marked preference for those aliphatic substrates which approach most closely the configurations of choline esters. The idea that pseudo-cholinesterases are less specific than true cholinesterases therefore receives no support from these studies. On the basis of their findings Adams and Whittaker proposed a classification of these enzymes according to the acyl group in the substrate molecule optimal for their activity, true cholinesterase being designated "aceto-cholinesterase" and the pseudo-type "butyro-cholinesterase". Unfortunately, when a wider range of animal species is studied, it becomes clear that the cholinesterases cannot be divided into only two types on a basis of substrate specificity, since Ord & Thompson (1951) have shown that the optimal acyl group for the cholinesterase of rat heart is propionyl, though it would seem from inhibitor studies that this enzyme is a pseudo-cholinesterase.

Despite the criticisms that have been levelled at Mendel's original classification, a case can still be made out for retaining the terms true and pseudo-cholinesterase, particularly if the criteria for classification are relative sensitivities to inhibition by selective inhibitors as well as substrate specificity.

4. Distribution of Cholinesterase Types in Different Tissues

In recent years considerable information has been obtained concerning the distribution and localization of true and pseudo-cholinesterases in different vertebrate tissues.

The erythrocytes of certain species and the brain had early been recognized as containing the true cholinesterase (Mendel & Rudney, 1943a; Nachmansohn & Rothenberg, 1945).

Pseudo-cholinesterase, which had been studied chiefly as a plasma enzyme, was next shown to be present in the pancreas (Mendel & Mundell, 1943); and Langemann (1944) found it also in human ovary, in contrast to human skeletal muscle, which contains only the true enzyme. The peripheral nerves of the cat, and the liver and certain glandular structures in the rat, were also shown to hydrolyse benzoylcholine (Sawyer & Hollinshead, 1945; Sawyer & Everett, 1947).

A more detailed study of the distribution of cholinesterase types in rat tissues was carried out by Ord & Thompson

(1950a), who were able to detect significant amounts of pseudo-cholinesterase in all the tissues studied except skeletal muscle. It appeared from their work that the rat tissues could be classified into three groups:

Group A. Tissues containing largely or entirely the true cholinesterase—brain, skeletal muscle, and the adrenal glands.

Group B. Tissues containing both enzymes in amounts such that acetyl-β-methylcholine and benzoylcholine are hydrolysed at about the same rate—stomach, liver, lung, and salivary gland.

Group C. Tissues containing chiefly pseudo-cholinesterase —heart, intestinal muscle and mucosa, Harderian gland, and skin.

The pseudo-cholinesterase contents of the heart (both auricle and ventricle) and intestine were strikingly high. In the rat therefore it would appear that those tissues in which acetylcholine exerts a nicotine-like action (i.e., in which it is concerned with transmission to another neurone or to a striated muscle cell) contain chiefly the true cholinesterase, while those in which it exerts a muscarine-like action (Brown, 1937) contain very largely the pseudo-cholinesterase; whether this apparent correlation is of physiological significance, however, is not yet known, and much more work is needed in other species before any final conclusion can be drawn.

Koelle (1950), using a histochemical technique, has studied the distribution in the cat, and has found that pseudo-cholinesterase predominates in the liver and in the ileum. That pseudo-cholinesterase may be actively secreted by glands is suggested by the work of McCance *et al.* (1951), who have claimed that dog pancreatic juice and pig parotid saliva are rich in this enzyme.

True cholinesterase, although most plentiful in nerve and striated muscle, is not confined to these tissues. It is of interest that it is the chief cholinesterase present in perfused, blood-free placenta (Ord & Thompson, 1950b), since this tissue is usually regarded as being devoid of nerves.

Turning to the nervous system, which Nachmansohn (1939) had studied in some detail as regards acetylcholine hydrolysis, it was concluded both by Mendel & Rudney (1943b) and by Nachmansohn & Rothenberg (1944) that brain contains only the true or "specific" cholinesterase. More recently it has been reported from a number of laboratories that brain does hydrolyse benzoylcholine, although at a relatively low rate (Augustinsson, 1948; Zeller, 1949; Ord & Thompson, 1950a). More detailed studies of different areas of the brain have been carried out by Burgen & Chipman (1951) and Ord & Thompson (1952), and it is now known that pseudo-cholinesterase is widely distributed in the central nervous system and that it appears to be particularly associated with the myelinated fibre tracts. In this connexion it is of interest that Koelle (1950) has shown the presence of pseudo-cholinesterase in glial cells and in the Schwann cells of myelinated nerves.

5. Effects of Inhibition of Cholinesterase

The effects produced in vivo by inhibition of the cholinesterases have been extensively studied in connexion with the pharmacology of the new organo-phosphorus anticholinesterases. To take DFP as an example of these compounds, the effects produced by its administration to animals can in the main be interpreted as resulting from an accumulation of acetylcholine due to inhibition of cholinesterase. Thus, lacrimation and profuse salivation occur; the sweat glands are stimulated, even after therapeutic doses of DFP (Grob *et al.* 1947b; Harvey *et al.* 1947); there is an increase in the tone and motility of the musculature of the gastrointestinal tract (Modell *et al.* 1946; Grob, Lilienthal & Harvey, 1947a); in the lungs there is constriction of bronchial muscles and increased secretion of bronchial glands (Green, McKay & Krop, 1947), while with large doses of DFP there is a progressive fall of blood pressure. In the eye an intense miosis is produced which may persist for weeks; application of DFP to the chronically denervated pupil, however, is not followed by constriction (Leopold & Comroe, 1946). In autonomic ganglia and at neuromuscular junctions, the facilitating effects which are produced are also in accord with an inhibition of cholinesterase at these sites.

It has been repeatedly observed that a very considerable inhibition, of the order of about 80%, of the true cholinesterase of the nervous system must be present before any abnormal signs develop which might be attributable to acetylcholine accumulation (Mazur & Bodansky, 1946; Koelle & Gilman, 1946; Nachmansohn & Feld, 1947). In the case of the pseudo-cholinesterase the level of this enzyme in the serum can be reduced almost to zero without the development of abnormal signs or symptoms (Mazur & Bodansky, 1946).

6. Physiological Functions of Cholinesterases

It is when we turn to considering the possible roles played by the cholinesterases inside the body that the suitability of the names true and pseudo-cholinesterase, in the light of our present knowledge concerning these enzymes, becomes apparent. For, although we have, on the whole, good evidence relating the true cholinesterase of synapses and neuromuscular junctions to the inactivation of physiologically released acetylcholine[1], the function of the pseudo-cholinesterase remains obscure, and much of the experimental evidence is interpreted by some workers in the field as arguing against any participation of this enzyme in acetylcholine metabolism, at any rate in the brain.

With regard to the true cholinesterase, Nachmansohn and his collaborators (Bullock *et al.* 1947) have put forward a more comprehensive theory of its function in transmission processes by suggesting that acetylcholine (and cholinesterase) is concerned also with conduction along the nerve fibre. This theory has, however, been extensively criticized.

It would certainly seem that, as pointed out above, no signs that could be referred to acetylcholine accumulation develop when the pseudo-cholinesterase of the plasma is virtually completely inactivated, and Hawkins & Gunter (1946) have claimed that this is so when the level of this enzyme in the pancreas and in the superior cervical ganglia is inhibited to the extent of 80–90%. On the other hand, it has recently been shown that certain late changes (demyelination of peripheral nerves and of certain tracts in the cord) can, under suitable conditions, be produced by certain selective inhibitors of pseudo-cholinesterase, including tri-*o*-cresyl phosphate, and that the level of the pseudo-cholinesterase in the spinal cord is profoundly lowered prior to the onset of the demyelination (Earl & Thompson, 1952). It cannot be

[1] See the review by Feldberg, W., in *Brit. med. Bull.* 1950, **6**, 312.

concluded, however, that the demyelination is the result of this enzyme inhibition. Although it is known that, in the hen, DFP and bis(*iso*propylamino)fluorophosphine oxide can cause demyelination, other organo-phosphorus anticholinesterases do not, even although considerable inhibition of pseudo-cholinesterase in the spinal cord is produced in animals poisoned by them (Thompson & Webster, unpublished). Further, there is evidence that, in the rat, inactivation of pseudo-cholinesterase by tri-*o*-cresyl phosphate in the central nervous system is not accompanied by any structural or functional changes (Mendel & Myers, 1952).

It has been proposed that pseudo-cholinesterase may act by supplementing the action of the true enzyme, since Koelle, Koelle & Friedenwald (1950) have found that concentrations of DFP which have little effect on the true cholinesterase, but which cause a marked fall in pseudo-cholinesterase activity, produce an increase in the tone and motility of the cat's ileum. Feldberg (1950) has concluded that the acetylcholine metabolism of the intestine is in part independent of nervous activity and has suggested that it may be responsible for stimulating the spontaneous contractions of the intestinal wall and the secretion of the succus entericus; it is possible that the pseudo-cholinesterase, which is present in such high concentrations in the wall of the intestine, may be concerned with such local actions.

Other observations, such as the rise in the level of pseudo-cholinesterase in the liver and in the plasma in animals fed on hypolipotropic diets, and the return to normal on adding supplements of lipotropic substances (Hawkins & Nishikawara, 1951), the fall in the level of the plasma cholinesterase in liver disease (Antopol, Tuchman & Schifrin, 1937; McArdle, 1940) and in malnutrition (McCance, Widdowson & Hutchinson, 1948; Waterlow, 1950; Hutchinson, McCance & Widdowson, 1951), together with the changes that have been reported in thyrotoxicosis (Antopol *et al.* 1937) and in alloxan diabetes (Saviano, Balbi & De Franciscis, 1948), cannot at present be interpreted with anything but the utmost reserve, although certain of them suggest that the pseudo-cholinesterase in liver might be concerned with some aspect of choline or lipoid metabolism.

References

Adams, D. H. & Thompson, R. H. S. (1948) *Biochem. J.* **42,** 170
Adams, D. H. & Whittaker, V. P. (1949a) *Biochim. biophys. Acta,* **3,** 1
Adams, D. H. & Whittaker, V. P. (1949b) *Biochim. biophys. Acta,* **3,** 358
Adrian, E. D., Feldberg, W. & Kilby, B. A. (1947) *Brit. J. Pharmacol.* **2,** 56
Aldridge, W. N. (1950) *Biochem. J.* **46,** 451
Alles, G. A. & Hawes, R. C. (1940) *J. biol. Chem.* **133,** 375
Antopol, W., Tuchman, L. & Schifrin, A. (1937) *Proc. Soc. exp. Biol., N.Y.* **36,** 46
Augustinsson, K.-B. (1948) *Acta physiol. scand.* **15,** Suppl. No. 52
Blaschko, H. & Holton, P. (1949) *Brit. J. Pharmacol.* **4,** 181
Bloch, H. (1941) *Helv. med. Acta,* **8,** Suppl. No. 7, p. 15
Brown, G. L. (1937) *Physiol. Rev.* **17,** 485
Bullock, T. H., Grundfest, H., Nachmansohn, D. & Rothenberg, M. A. (1947) *J. Neurophysiol.* **10,** 11
Burgen, A. S. V. & Chipman, L. M. (1951) *J. Physiol.* **114,** 296
Dale, H. H. (1914) *J. Pharmacol.* **6,** 147
Earl, C. J. & Thompson, R. H. S. (1952) *Brit. J. Pharmacol.* **7,** 261
Easson, L. H. & Stedman, E. (1937) *Biochem. J.* **31,** 1723
Englehart, E. & Loewi, A. (1930) *Arch. exp. Path. Pharmak.* **150,** 1
Feldberg, W. S. (1950) *Proc. roy. Soc.* B, **137,** 285
Glick, D. (1941) *Nature, Lond.* **148,** 662
Green, R. E., McKay, E. A. & Krop, S. (1947) *Fed. Proc.* **6,** 334
Grob, D., Lilienthal, J. L. & Harvey, A. M. (1947a) *Johns Hopk. Hosp. Bull.* **81,** 245
Grob, D., Lilienthal, J. L., Harvey, A. M. & Jones, B. F. (1947b) *Johns Hopk. Hosp. Bull.* **81,** 217
Harvey, A. M., Lilienthal, J. L., Grob, D., Jones, B. F. & Talbot, S. A. (1947) *Johns Hopk. Hosp. Bull.* **81,** 237
Hawkins, R. D. & Gunter, J. M. (1946) *Biochem. J.* **40,** 192
Hawkins, R. D. & Mendel, B. (1947) *Brit. J. Pharmacol.* **2,** 173
Hawkins, R. D. & Nishikawara, M. T. (1951) *Biochem. J.* **48,** 276
Holton, P. (1948) *Biochem. J.* **43,** *Proc.* xiii
Hutchinson, A. O., McCance, R. A. & Widdowson, E. M. (1951) *Spec. Rep. Ser. med. Res. Coun., Lond.* No. 275, p. 216
Koelle, G. B. (1950) *J. Pharmacol.* **100,** 158
Koelle, G. B. & Gilman, A. (1946) *J. Pharmacol.* **87,** 421
Koelle, G. B., Koelle, E. S. & Friedenwald, J. S. (1950) *J. Pharmacol.* **100,** 180
Langemann, K. (1944) *Helv. physiol. Acta,* **2,** C17
Leopold, I. H. & Comroe, J. H. (1946) *Arch. Ophthal., N.Y.* **36,** 17
Levine, M. G. & Suran, A. A. (1950) *Nature, Lond.* **166,** 698
Loewi, O. & Navratil, E. (1926) *Pflüg. Arch. ges. Physiol.* **214,** 678
McArdle, B. (1940) *Quart. J. Med.* **33,** 107
McCance, R. A., Brown, L. M., Comline, R. S. & Titchen, D. A. (1951) *Nature, Lond.* **168,** 788
McCance, R. A., Widdowson, E. M. & Hutchinson, A. O. (1948) *Nature, Lond.* **161,** 56
Mackworth, J. F. & Webb, E. C. (1948) *Biochem. J.* **42,** 91
Matthes, K. (1930) *J. Physiol.* **70,** 338
Mazur, A. & Bodansky, O. (1946) *J. biol. Chem.* **163,** 261
Mendel, B. & Mundell, D. B. (1943) *Biochem. J.* **37,** 64
Mendel, B. & Myers, D. K. (1952) *Nature, Lond.* **170,** 928
Mendel, B. & Rudney, H. (1943a) *Biochem. J.* **37,** 59
Mendel, B. & Rudney, H. (1943b) *Science,* **98,** 201
Mendel, B. & Rudney, H. (1944) *Science,* **100,** 499
Modell, W., Krop, S., Hitchcock, P. & Riker, W. F., jr (1946) *J. Pharmacol.* **87,** 400
Nachmansohn, D. (1939) *Bull. Soc. Chim. biol., Paris,* **21,** 761
Nachmansohn, D. & Feld, E. A. (1947) *J. biol. Chem.* **171,** 715
Nachmansohn, D. & Rothenberg, M. A. (1944) *Science,* **100,** 454
Nachmansohn, D. & Rothenberg, M. A. (1945) *J. biol. Chem.* **158,** 653
Ord, M. G. & Thompson, R. H. S. (1950a) *Biochem. J.* **46,** 346
Ord, M. G. & Thompson, R. H. S. (1950b) *Nature, Lond.* **165,** 927
Ord, M. G. & Thompson, R. H. S. (1951) *Biochem. J.* **49,** 191
Ord, M. G. & Thompson, R. H. S. (1952) *Biochem. J.* **51,** 245
Richter, D. & Croft, P. G. (1942) *Biochem. J.* **36,** 746
Saviano, M., Balbi, R. & De Franciscis, P. (1948) *Boll. Soc. ital. Biol. sper.* **24,** 1353
Sawyer, C. H. & Everett, J. W. (1947) *Amer. J. Physiol.* **148,** 675
Sawyer, C. H. & Hollinshead, H. (1945) *J. Neurophysiol.* **8,** 137
Stedman, E., Stedman, E. & Easson, L. H. (1932) *Biochem. J.* **26,** 2056
Thompson, R. H. S. (1947) *J. Physiol.* **105,** 370
Waterlow, J. (1950) *Lancet,* **1,** 908
Whittaker, V. P. (1951) *Physiol. Rev.* **31,** 312
Zeller, E. A. (1949) *Helv. chim. Acta,* **32,** 484
Zeller, E. A. & Bissegger, A. (1943) *Helv. chim. Acta,* **26,** 619

ACTION OF DRUGS ON ENZYME SYSTEMS

J. H. QUASTEL D.Sc. F.R.S.

Department of Biochemistry
McGill University, Montreal, Canada

1 Competitive inhibition
2 Arsenicals
3 Narcotics
4 Reactions of acetylcholine
References

This subject has now assumed such large dimensions that it is impossible, in this short statement, to give more than a brief survey, and attention is paid to those aspects with which the writer has been associated.

1. Competitive Inhibition

The principle of competitive inhibition is widely accepted today as an aid to the understanding of the mechanism of action of many drugs and as a means of investigation into the problems of chemotherapy. Many reviews have been written on the subject (e.g., Fildes *et al.* 1949; Woolley, 1952; Roblin, 1949; Potter & Heidelberger, 1950).

The modern conception of the principle of competitive inhibition by structural analogues arose with the work of Quastel & Wooldridge (1928) on dehydrogenase systems of bacteria. They concluded that the active centre of an enzyme is so constituted that it may combine with a variety of substances all possessing a particular type of chemical structure, but that, of the total number of molecules of this structure, only a few are substrates capable of being activated to undergo subsequent chemical change. It followed from this view that many molecules may compete with each other and with the substrate for attachment to an enzyme. Thus, the presence of the structural analogues of the substrate of the enzyme would have the effect of diminishing the activity of the enzyme toward the substrate.

An important example of this phenomenon is the inhibitive effect of malonic acid ($COOH.CH_2.COOH$) on succinic acid ($COOH.CH_2.CH_2COOH$) dehydrogenase. Quastel & Wooldridge (1928) found that the chemical structure necessary for combination with succinic acid dehydrogenase seemed to be $—C.CH_2COOH$ or $—C.CH.COOH$. Similarly, substances having the structure .CO.COOH or .CHOH.COOH seemed to be common to those compounds that combined with lactic dehydrogenase. The type of molecule which will combine with one dehydrogenase may have little or no affinity for the other dehydrogenases.

It was evident that structural analogues, which would combine with an enzyme but which would not be activated, would act as competitive antagonists to the substrate if they were present in the enzyme system at sufficiently high concentrations, the extents of inhibition being dependent on the relative affinities of analogues and of substrates to the enzyme.

That a structural analogue may also act as an antimetabolite in the modern sense of the word was early demonstrated in experiments by Quastel & Wooldridge (1929) on the conditions underlying the proliferation of *Bacterium coli* in synthetic media. They showed that the proliferation of these organisms in a medium where the sole source of carbon was provided by succinate was greatly inhibited if malonate was also present in the medium. Although the malonate acted as an inhibitor of bacterial growth in this system it was not a cell poison, for it had no inhibitory effects in systems where the preliminary oxidation of succinate was not necessary to provide the bacteria with carbon for synthesis.

The inhibitory effect of malonate on succinic dehydrogenase has been of the greatest importance in the study of respiratory systems of the cell. Malonate inhibits succinate oxidation in mammalian tissues (Quastel & Wheatley, 1931) and acetoacetate breakdown in liver and kidney tissues (Jowett & Quastel, 1935). The phenomenon of malonate inhibition was fundamental for the studies of Szent-Györgyi (1936) on fumaric acid catalysis in muscle respiration and for the later work of Krebs (1943) and other investigators into the role of the tricarboxylic acid cycle in metabolism.

It must be borne in mind, however, that malonate has structural similarity to substances other than succinate and, therefore, might compete as a structural analogue with substrates of enzymes other than succinic dehydrogenase. It is, for example, an inhibitor of the decarboxylation of oxaloacetate by liver (Evans, Vennesland & Slotin, 1943; Salles & Ochoa, 1950).

Modern interest in the principle of competitive inhibition commenced when Woods (1940)[1] reported that the bacteriostatic action of sulphanilamide and related sulphonamide drugs was reversed competitively by *p*-aminobenzoic acid. The competitive relationship, taken together with a similarity in chemical structure between the antagonist and sulphanilamide and the high activity of *p*-aminobenzoic acid, led to a working hypothesis concerning the mode of action of sulphonamides. It was suggested that *p*-aminobenzoate is an essential substance for the growth of bacteria and that sulphonamides, by virtue of their chemical similarity, inhibit competitively the enzyme system involved in the utilization of this substrate. The hypothesis received its most important confirmation by the discovery that *p*-aminobenzoate is in fact an essential growth factor of bacteria. There is now evidence to indicate that *p*-aminobenzoate is a precursor directly or indirectly in the biosynthesis of folic acid, purines, thymine, and certain amino-acids, for example, methionine, serine and glycine.

Recognition of the importance of the principle of competitive inhibition as a possible explanation of the mode of action of the sulphanilamides stimulated a search for analogues of other known bacterial growth factors. This concept, made clear by Fildes (1940), resulted in the discovery of a large number of structural analogues that inhibit specifically the growth of micro-organisms.

Nicotinic acid analogues, for example 3-pyridinesulphonic acid, and pantothenic acid analogues, such as pantoyltaurine, act competitively in producing inhibitions of bacterial proliferation (McIlwain, 1940, 1944, 1945; McIlwain & Hughes, 1944; Snell, 1941; Shive & Snell, 1945; Woolley & Collyer, 1945). Further advance was made by Woolley & White (1943) when they found that certain types of structural

[1] See also Woods, *Brit. med. Bull.* 1948, 5, 308; and this Bulletin, p. 122.

analogues of various vitamins would cause the appearance of characteristic signs of vitamin deficiency diseases in animals. These signs were prevented by adequate dosage with the vitamin concerned. For example, an analogue of thiamine, pyrithiamine, will produce the symptoms characteristic of thiamine deficiency in mice, and the disease may be prevented or cured by adding sufficient amounts of thiamine. Woolley & Krampitz (1943) showed that glucoascorbic acid, a structural analogue of ascorbic acid, produces a scurvy-like disease of rats and mice and guinea-pigs, and that in guinea-pigs the disease is prevented by the addition of ascorbic acid. Riboflavin deficiency may be produced in rats by feeding *iso*riboflavin and in mice by feeding the phenazine analogue of riboflavin. Production of nicotinic acid deficiency may be brought about in mice by feeding free acetylpyridine and this may be prevented by administration of nicotinic acid.

2:4-Diaminopyrimidines and other substances containing the 2:4-diaminopyrimidine group are competitive antagonists of folic acid in the growth of bacteria, e.g., *Lactobacillus casei*, *Streptococcus faecalis* (Hitchings *et al.* 1952). Various other antagonists of this type (e.g., aminopterin) are inhibitory to bacterial growth (Jukes, Franklin & Stokstad, 1950) and such substances have been used in investigations on leukaemia. Hutchison & Burchenal (1952) point out that cells resistant to these substances may be capable of utilizing the antagonist or may have developed enzymes less vulnerable to the antagonist.

Benzimidazole, an agent that causes in animals loss of muscular tone and of ability to respond to stimuli, prevents the growth of many micro-organisms. Its growth-inhibiting action is reversed competitively by adenine. Adenine, however, does not reverse the effects of benzimidazole in animals. It may well be that benzimidazole affects various enzyme systems, its structure having features in common with those of the substrates of different enzymes. This is exemplified by recent observations on the behaviour of benzedrine (phenyl-*iso*propylamine) with enzyme systems.

Mann & Quastel (1940) showed that benzedrine and ephedrine are capable of competitively inhibiting amine oxidase of rat brain (Pugh & Quastel, 1937), and suggested that this effect may be intimately concerned with the effect of benzedrine on the central nervous system. Oota (1946), using more than 30 benzedrine derivatives, has extended these observations. He found a fairly close parallelism between the ability to inhibit amine oxidase and the stimulating effect on the central nervous system. Colter & Quastel (1952) have now found that benzedrine, ephedrine and various other amines are able to compete with choline for choline oxidase, the affinity of benzedrine for choline oxidase being even higher than it is for amine oxidase. It is evident that benzedrine may affect several enzyme systems and its pharmacological effects may be dependent on its affinity for more than one enzyme.

Interest is being taken today in the mode of action of methionine sulphoximine, a causative agent of canine hysteria. This substance, elaborated in the process of manufacturing certain foods, produces in dogs convulsive seizures (Mellanby, 1946). It is stated that the convulsive action, in dogs or in rabbits, of this analogue of methionine can be overcome by the administration of methionine. Methionine sulphoxide, which is a structural analogue of glutamic acid, inhibits the conversion of glutamate to glutamine; apparently the toxicity of methionine sulphoxide is related to this phenomenon (Borek *et al.* 1946; Waelsch *et al.* 1946; McIlwain, Roper & Hughes, 1948; Roper & McIlwain, 1948; Elliott, 1951) Braganca, Faulkner & Quastel (1953) have now found that methionine sulphoxide and methionine sulphoximine, both of which are inhibitors of glutamine synthesis, affect the rate of synthesis of acetylcholine in the nervous system, by inhibiting the utilization of adenosinetriphosphate required in glutamine synthesis. This is an example of a structural analogue influencing a metabolic reaction indirectly.

New examples of competitive inhibition of metabolic processes by structural analogues are being continually discovered, providing important material in the search for new drugs.

2. Arsenicals

Arsenicals are examples of drugs that act by non-competitive combination with a specific group of the active centre of an enzyme. It is known (Ehrlich, 1909) that while quinquevalent arsenic components are relatively non-toxic to protozoa in vitro, tervalent arsenic compounds such as the substituted phenylarsenoxides are extremely toxic. It seems that this toxicity is due to the ability of the arsenoxides to react with thiol compounds (Voegtlin, Dyer & Leonard, 1923). The fact that arsenic compounds (phenylarsenoxides) are capable of combining with organic compounds containing SH groups was demonstrated by Barber (1929), and Gough & King (1930) concluded that, when quinquevalent phenylarsonic acid derivatives are given therapeutically, reduction takes place in the tissue to give arsenoxides which then condense with thiol compounds of the tissues.

Yorke, Adams & Murgatroyd (1929) showed that glucose consumption and respiration of pathogenic trypanosomes were extinguished simultaneously by trypanocidal arsenoxides and concluded that trypanocidal drugs act by inhibiting glucose metabolism.

Following an investigation by Hawking, Hennelly & Quastel (1937), who had found that administration of the quinquevalent arsenical tryparsamide into the body led to the appearance of tervalent arsenic compounds of highly trypanocidal activity in the spinal fluid, Gordon & Quastel (1947, 1948) carried out an investigation to discover the nature of the combination of organic tervalent and quinquevalent arsenic compounds with tissue proteins and enzyme systems. This work, carried out during 1938 and 1939 and reported to the Medical Research Council in 1940 (see Gordon & Quastel, 1947), led to the conclusion that reversible equilibria exist between tervalent arsenic compounds (such as derivatives of phenylarsenoxide) and purified tissue proteins, and that certain compounds of the type *R*.AsO, e.g., mapharside and phenylarsenoxide, combine reversibly with enzymes characterized by the possession of SH groups. Such enzymes include urease, succinic dehydrogenase, and choline dehydrogenase. Enzymes known not to possess thiol groups are unaffected by the phenylarsenoxides. The toxicity of the phenylarsenoxides to enzymes is neutralized by the addition of excess thiol compounds, e.g., cysteine. These condense with the phenylarsenoxides to form easily dissociated compounds, and excess thiol compound is required to keep most of the *R*.AsO in the form of an inert complex. Similar results have recently been obtained with carboxylase (Stoppani *et al.* 1952). Compounds of the type *R*.AsO and *R*.$AsCl_2$ also inhibit the respiration of intact cells in the presence of glucose and pyruvate. The inhibition of pyruvic oxidase by mapharside (Gordon & Quastel, 1947, 1948) is an indication of the SH nature of

this enzyme; the inhibition is reversed by the presence of excess thiol compounds. Work independently carried out by Barron & Singer (1945), and by Dixon & Needham (1946), has demonstrated the sensitivity of SH enzymes to lewisite and various arsenoxides. Peters (1948, 1952), Stocken & Thompson (1949) and Peters, Stocken & Thompson (1945) have shown that the toxic effects of lewisite are due to a combination of the mono-substituted arsenic atom with two adjacent thiol groups in a protein constituent of the pyruvate oxidase system.[2] A ring compound is formed which has considerable stability and which cannot be reversed by a monothiol compound such as glutathione. It is reversed, however, by transfer of the arsenic grouping to British anti-lewisite, 2:3-dimercapto-propanol, because the 5-membered ring formed by this substance is more stable than that with a tissue component. The advent of British anti-lewisite (BAL) through the work of Peters and his colleagues has resulted in a substance of considerable therapeutic importance. This dithiol not only combines with toxic arsenic compounds but also with mercury and lead compounds, which exert toxicity probably by combination with the thiol groups of tissue enzymes.

3. Narcotics[3]

If various alkyl barbiturates are compared as to their inhibitory effects on the oxygen consumption of minced guinea-pig brain, it is found (Quastel & Wheatley, 1932) that there exists a parallelism between their hypnotic and inhibitive powers, this parallelism being shown among narcotics of different chemical types (see also Fuhrman & Field, 1943). Definite inhibition of respiration (15–32%) of cerebral cortical slices occurs in the presence of narcotics (barbiturates, chloretone) at concentrations that produce narcosis in animals (Jowett & Quastel, 1937a, 1937b; Jowett, 1938). The inhibitions recorded represent the over-all effects of narcotics on the respiration of the entire cerebral cortex of the animal; they may be greater at those parts of the nervous system where the narcotic is localized or specifically absorbed or for which the narcotic may have a high affinity.

Narcotics do not inhibit all oxidative processes to the same extent. The oxidation of glucose, lactate or pyruvate is most affected, that of glutamate is affected to a less extent, whilst that of succinate or of *p*-phenylenediamine is undisturbed.

The high sensitivity of glucose, or of pyruvate, oxidation in the brain to such narcotics as the barbiturates or chloretone is a striking feature of narcotic action in vitro and, in view of the great importance of glucose or pyruvate in the functional activity of the central nervous system, this sensitivity seems to be a highly important factor in any consideration of the mechanism of narcotic action.

The inhibitive effects of narcotics, such as the barbiturates or chloretone, on the respiration of brain slices are reversible (Quastel & Wheatley, 1934). High concentrations of narcotics, however, produce irreversible effects. Analysis of the kinetic data indicates that two effects of a narcotic on brain respiration in vitro take place. There is a rapid attainment of an equilibrium between the narcotic and a constituent of the respiratory system. This follows the mass action law so long as the inhibition does not exceed 40%, and applies to urethane, chloretone, chloral, barbiturates, avertin and magnesium ions. There is also a slow development of irreversible changes leading to increased inhibition of respiration that cannot be restored to normal by removal of the narcotic. This takes place with most narcotics but it is observable only at relatively high concentrations with narcotics such as barbiturates or chloretone. It occurs, however, at low concentrations with ether or ethanol. Conceivably this is due to gradual irreversible denaturation of the enzyme with which the drug becomes associated.

The steady rate of the diminished respiration of brain slices brought about by small concentrations of luminal or chloretone is greatly dependent upon the concentration of potassium ions. At low concentrations of potassium in the medium, the respiration, in presence of narcotic, falls progressively, this being probably due to a leak of potassium from the cell at low external concentration (Jowett & Quastel, 1937a). It is known that the presence of potassium is important for metabolic activities of the nerve cell.

Elliott, Warrens & James (1947) have examined the effects of demerol and amidone on brain oxidations. These substances exert powerful effects at relatively high concentrations on lactate and pyruvate oxidation.

It is also known that steroids such as deoxycorticosterone and progesterone, that have anaesthetic potency, also affect the respiration of rat brain homogenates in presence of glucose but not in presence of succinate. The inhibitions parallel anaesthetic action (Gordan & Elliott, 1947).

Low concentrations of narcotics do not affect the known dehydrogenases. There is no inhibition, for example, of the dehydrogenase systems involved in anaerobic glycolysis. There is indeed evidence to show that narcotics increase aerobic glycolysis and glucose breakdown by suppression of the Pasteur reaction (Rosenberg *et al.* 1950; Webb & Elliott, 1951).

Analysis of the various links in the respiratory chain leading to glucose oxidation has indicated that there is a factor highly sensitive to narcotics in this multi-enzyme system (Michaelis & Quastel, 1941; Greig, 1946).

Recent evidence by Johnson & Quastel (1953) leads to the conclusion that narcotics, such as chloretone or nembutal, at low concentrations inhibit an oxidative step in cell respiration concerned with the synthesis of adenosinetriphosphate, and thereby inhibit biological acetylations such as those of choline or of sulphanilamide. The addition of adenosinetriphosphate to the aerobic systems relieves the inhibition of acetylation brought about by the narcotic. The narcotics even at relatively high concentrations do not inhibit the anaerobic acetylations of sulphanilamide or the anaerobic synthesis of acetylcholine. Eiler & McEwen (1949) had already shown that pentabarbital inhibits the generation of high energy phosphate bonds to the extent that it interferes with oxygen utilization (see also McElroy, 1947). Brody & Bain (1951) have, indeed, postulated that uncoupling of phosphorylation from oxidation may be connected with the narcotic action of barbiturates. The effects of narcotics on phosphorylations differ markedly, however, from those of a typical uncoupling agent such as 2:4-dinitrophenol (Johnson & Quastel, 1953).

Against the view that a narcotic exercises its activity by a suppression, even a specific one, of oxidative events at the nerve centre, where it is taken up, are the results of the experiments of Larrabee (1952) to the effect that the oxygen uptake shown by isolated nerve does not cease when transmission of an impulse is blocked by the presence of a narcotic such as chloretone. If the nerve is stimulated there is an increased oxygen consumption and in the presence of a

[1] See Peters, p. 116 of this Bulletin.
[2] For reviews of biochemical aspects see Quastel (1952) and Harris (1951).

narcotic this oxygen consumption is greatly depressed (Bronk & Brink, 1951). The lack of any measure of retardation of the rate of oxygen consumption in a resting nerve by a narcotic is not necessarily opposed to the conclusions that have already been made. In a multi-enzyme system, such as that present in a chain of respiratory events, it is possible that one step, important for the functional activity of the cell, may be impeded more than the succeeding steps. The total activity of the respiratory chain may not be measurably affected so long as the enzymes following the impeded step are still saturated with their respective substrates. An example of this is the small, or apparent lack of, inhibitive effect of malonate on the oxygen consumption of bacteria respiring in the presence of succinate. This is due to the fact that the remaining activity of the succinic dehydrogenase is still sufficient to supply the succeeding enzymes with the substances they require for their optimal activities (Quastel & Wheatley, 1931).

4. Reactions of Acetylcholine

Two classes of drugs seem to affect reactions undergone by acetylcholine:

i. Those compounds which inhibit cholinesterase and whose presence, therefore, results in an increased quantity of acetylcholine being available in the nervous tissue. These drugs are the anticholinesterases, whose effects on the nervous system are well known and on which there is a large literature.[4] Such substances are now becoming important as insecticides (Aldridge & Barnes, 1952; Myers *et al.* 1952).

ii. Those substances that compete with acetylcholine for receptor proteins in the nervous, or neuromuscular, system. Examples are paludrine whose application can produce arrest of the auricles of the heart (Burn, 1950) and curare-like substances (Grundfest, Nachmansohn & Chambers, 1952). It has already been pointed out (Quastel & Tennenbaum, 1937) that the narcotics morphine, heroin and codeine also appear to compete with acetylcholine for certain receptor proteins. It seems possible that the quaternary ammonium bases, which exert, at low concentrations, but little effect on respiration, act by competition with acetylcholine, rather than by an interference with a metabolic process.

[4] See Thompson, p. 138 of this Bulletin.

REFERENCES

Aldridge, W. N. & Barnes, J. M. (1952) *Nature, Lond.* **169,** 345
Barber, H. J. (1929) *J. chem. Soc.*, p. 1020
Barron, E. S. G. & Singer, T. P. (1945) *J. biol. Chem.* **157,** 221
Borek, E., Miller, H. K., Scheiness, P. & Waelsch, H. (1946) *J. biol. Chem.* **163,** 347
Braganca, B. M., Faulkner, P. & Quastel, J. H. (1953) *Biochim. biophys. Acta,* **10,** 83
Brody, T. M. & Bain, J. A. (1951) *Proc. Soc. exp. Biol., N.Y.* **77,** 50
Bronk, D. W. & Brink, F. (1951) *Fed. Proc.* **10,** 19
Burn, J. H. (1950) *Proc. roy. Soc.* B, **137,** 281
Colter, J. S. & Quastel, J. H. (1952) *Arch. Biochem. Biophys.* **41,** 305
Dixon, M. & Needham, D. M. (1946) *Nature, Lond.* **158,** 432
Ehrlich, P. (1909) *Ber. dtsch. chem. Ges.* **42,** 42
Eiler, J. J. & McEwen, W. K. (1949) *Arch. Biochem.* **20,** 163
Elliott, H. W., Warrens, A. E. & James, H. P. (1947) *J. Pharmacol.* **91,** 98
Elliott, W. H. (1951) *Biochem. J.* **49,** 106
Evans, E. A., Vennesland, B. & Slotin, L. (1943) *J. biol. Chem.* **147,** 771
Fildes, P. (1940) *Lancet,* **1,** 955
Fildes, P., Woods, D. D., McIlwain, H., Work, T. S., Rydon, H. N. & Rose, F. L. (1949) *Proc. roy. Soc.* B, **136,** 147
Fuhrman, F. A. & Field, J. (1943) *J. Pharmacol.* **77,** 392
Gordan, G. S. & Elliott, H. W. (1947) *Endocrinology,* **41,** 517
Gordon, J. J. & Quastel, J. H. (1947) *Nature, Lond.* **159,** 97
Gordon, J. J. & Quastel, J. H. (1948) *Biochem. J.* **42,** 337
Gough, G. A. C. & King, H. (1930) *J. chem. Soc.*, p. 669
Greig, M. C. (1946) *J. Pharmacol.* **87,** 185
Grundfest, H., Nachmansohn, D. & Chambers, R. (1952) *Nature, Lond.* **169,** 190
Harris, T. A. B. (1951) *The mode of action of anaesthetics.* Livingstone, Edinburgh
Hawking, F., Hennelly, T. J. & Quastel, J. H. (1937) *J. Pharmacol.* **59,** 157
Hitchings, G. H., Falco, E. A., VanderWerff, H., Russell, P. B. & Elion, G. B. (1952) *J. biol. Chem.* **199,** 43
Hutchison, D. J. & Burchenal, J. H. (1952) *Proc. Soc. exp. Biol., N.Y.* **81,** 251
Johnson, W. J. & Quastel, J. H. (1953) *Nature, Lond.* **171,** 602
Jowett, M. (1938) *J. Physiol.* **92,** 322
Jowett, M. & Quastel, J. H. (1935) *Biochem. J.* **29,** 2186
Jowett, M. & Quastel, J. H. (1937a) *Biochem. J.* **31,** 565
Jowett, M. & Quastel, J. H. (1937b) *Biochem. J.* **31,** 1151
Jukes, T. H., Franklin, A. L. & Stokstad, E. L. R. (1950) *Ann. N.Y. Acad. Sci.* **52,** 1336
Krebs, H. A. (1943) *Advanc. Enzymol.* **3,** 191
Larrabee, M. G. (1952) *Biology of mental health and disease,* p. 384. Hoeber, New York
McElroy, W. D. (1947) *Quart. Rev. Biol.* **22,** 25
McIlwain, H. (1940) *Brit. J. exp. Path.* **21,** 136
McIlwain, H. (1944) *Biochem. J.* **38,** 97
McIlwain, H. (1945) *Biochem. J.* **39,** 329
McIlwain, H. & Hughes, D. E. (1944) *Biochem. J.* **38,** 187
McIlwain, H., Roper, J. A. & Hughes, D. E. (1948) *Biochem. J.* **42,** 492
Mann, P. J. G. & Quastel, J. H. (1940) *Biochem. J.* **34,** 414
Mellanby, E. (1946) *Brit. med. J.* **2,** 885
Michaelis, M. & Quastel, J. H. (1941) *Biochem. J.* **35,** 518
Myers, D. K., Mendel, B., Gersmann, H. R. & Ketelaar, J. A. A. (1952) *Nature, Lond.* **170,** 805
Oota, Y. (1946) *J. chem. Soc. Japan,* **67,** 45
Peters, R. A. (1948) *Brit. med. Bull.* **5,** 313
Peters, R. A. (1952) *Proc. roy. Soc.* B, **139,** 143
Peters, R. A., Stocken, L. A. & Thompson, R. H. S. (1945) *Nature, Lond.* **156,** 616
Potter, V. R. & Heidelberger, C. (1950) *Physiol. Rev.* **30,** 487
Pugh, C. E. M. & Quastel, J. H. (1937) *Biochem. J.* **31,** 2306
Quastel, J. H. (1952) *Curr. Res. Anesth.* **31,** 151
Quastel, J. H. & Tennenbaum, M. (1937) *J. Pharmacol.* **60,** 228
Quastel, J. H. & Wheatley, A. H. M. (1931) *Biochem. J.* **25,** 117
Quastel, J. H. & Wheatley, A. H. M. (1932) *Proc. roy. Soc.* B, **112,** 60
Quastel, J. H. & Wheatley, A. H. M. (1934) *Biochem. J.* **28,** 1521
Quastel, J. H. & Wooldridge, W. R. (1928) *Biochem. J.* **22,** 689
Quastel, J. H. & Wooldridge, W. R. (1929) *Biochem. J.* **23,** 115
Roblin, R. O. (1949) *Chem. Engng News,* **27,** 3624
Roper, J. A. & McIlwain, H. (1948) *Biochem. J.* **42,** 485
Rosenberg, A. J., Buchel, L., Etling, N. & Levy, J. (1950) *C. R. Acad. Sci., Paris,* **230,** 480
Salles, J. P. V. & Ochoa, S. (1950) *J. biol. Chem.* **187,** 849
Shive, W. & Snell, E. E. (1945) *Science,* **102,** 401
Snell, E. E. (1941) *J. biol. Chem.* **139,** 975; **141,** 121
Stocken, L. A. & Thompson, R. H. S. (1949) *Physiol. Rev.* **29,** 168
Stoppani, A. O. M., Actis, A. S., Deferrari, J. O. & Gonsalez, E. L. (1952) *Nature, Lond.* **170,** 842
Szent-Györgyi, A. (1936) *Hoppe-Seyl. Z.* **244,** 105
Voegtlin, C., Dyer, H. & Leonard, C. S. (1923) *Publ. Hlth Rep., Wash.* **38,** 1882
Waelsch, H., Owades, P., Miller, H. K. & Borek, E. (1946) *J. biol. Chem.* **166,** 273
Webb, J. L. & Elliott, K. A. C. (1951) *J. Pharmacol.* **103,** 24
Woods, D. D. (1940) *Brit. J. exp. Path.* **21,** 74
Woolley, D. W. (1952) *A study of antimetabolites.* Wiley, New York
Woolley, D. W. & Collyer, M. L. (1945) *J. biol. Chem.* **159,** 263
Woolley, D. W. & Krampitz, L. O. (1943) *J. exp. Med.* **78,** 333
Woolley, D. W. & White, A. G. C. (1943) *J. biol. Chem.* **149,** 285
Yorke, W., Adams, H. R. D. & Murgatroyd, F. (1929) *Ann. trop. Med. Parasit.* **23,** 501

ENZYMIC OXIDATION OF AMINES

H. BLASCHKO M.D.

Department of Pharmacology
University of Oxford

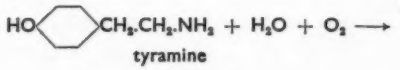
1 Discovery of amine oxidase
2 Properties of amine oxidase
3 Substrate specificity of amine oxidase
4 Histaminase and amine oxidase
5 Competitive inhibitors of amine oxidase
6 Distribution of amine oxidase
7 Amine oxidase and amine metabolism in the living organism
8 Physiological significance of amine oxidase
9 Derivatives of tryptamine
10 Amine oxidase in invertebrates
References

Until about twenty years ago, most of our knowledge of the metabolism of amines in the organism came from feeding experiments on animals and from observations on perfused organs. Amines not normally found in the body were administered and characteristic breakdown products were isolated from the urine or from perfusion fluids. These products were characterized by one common feature: the amino nitrogen had been split off and was excreted as urea.

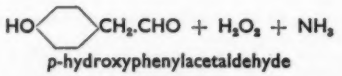
This type of amine metabolism was considered as a typical example of a "detoxication mechanism". We now know that "detoxication" of a synthetic compound usually indicates that related substances normally occurring in the body are broken down by an analogous route. These normal metabolites are often split to very simple end-products which do not reveal the pathway of the degradation reaction. The study of a "detoxication mechanism" has therefore a twofold interest: it tells us something about the fate of synthetic substances, e.g., of compounds of therapeutic interest, but it also causes us to search for substances normally present in the body that are broken down by the same metabolic route.

1. Discovery of Amine Oxidase

In 1928 Miss M. L. C. Hare (now Mrs Bernheim), at the Cambridge School of Biochemistry, demonstrated the presence in the mammalian liver of an enzyme that catalysed the oxidative deamination of tyramine. This reaction can be formulated thus:

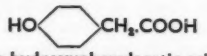


tyramine



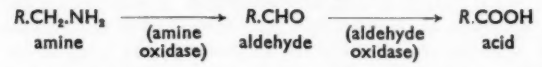
p-hydroxyphenylacetaldehyde

The fate of tyramine in the intact organism and in perfused organs had already been studied by Ewins & Laidlaw (1910); in their experiments, the end-product of tyramine breakdown was *p*-hydroxyphenylacetic acid:



p-hydroxyphenylacetic acid

The enzyme discovered by Hare accounted for the metabolic fate of tyramine: in the organism, the aldehyde formed as the first oxidation product would be expected to be further oxidized to the corresponding carboxylic acid:



Although tyramine may be formed as a product of the decarboxylation of L-tyrosine, it is not normally present in the mammalian organism in significant amounts. The work on amine metabolism therefore led to two parallel lines of inquiry: firstly, to the study of the substrate specificity of the enzyme discovered by Hare, and, secondly, to attempts to elucidate its significance in normal cell metabolism.

The concept of amine oxidase, as it is studied today, arose out of work carried out in Cambridge and Cardiff in the years before the war (Blaschko, Richter & Schlossmann, 1937a, 1937b; Kohn, 1937; Pugh & Quastel, 1937a, 1937b). The enzyme that oxidized tyramine was found to be widely distributed in mammalian tissues; it acted not only on tyramine, but on very many other sympathomimetic amines, including adrenaline and noradrenaline; many simple aliphatic amines were also oxidized. Tryptamine was another compound of biological interest that was oxidized by amine oxidase.

That the oxidation of all these amines is catalysed by one enzyme is supported (i) by the study of the distribution of the deamination reaction and (ii) by competition experiments: (i) as far as is known, tissues with amine oxidase activity will act on tyramine as well as on aliphatic amines, and (ii) when one amine is present in saturation concentration, the addition of a second amine does not increase the oxygen consumption; in other words, the two amines compete for the enzyme. However, amine oxidase has never been purified or isolated, and it is possible that several closely related enzymes exist which usually occur together. This is the view of Werle & Roewer (1952), who believe that, in the ox liver, tyramine and *n*-butylamine are oxidized by two different enzymes.

2. Properties of Amine Oxidase

Amine oxidase is usually studied in crude tissue suspensions or homogenates, but the enzymic activity is fully retained in acetone-dried preparations. In the liver cell, the enzyme is present in the cytoplasmatic granules (Cotzias & Dole, 1951; Hawkins, 1952). Hawkins found most of the activity in the mitochondria, but some activity was also present in the smaller particles, the so-called microsomes; it is not yet known if these particles were all preformed cell structures or if they included fragmented mitochondria.

The mitochondria contain the respiratory catalysts of the cell, and the fact that amine oxidase is also found in these elements suggests that the oxidation of amines is bound up with the other respiratory activities of the cell. This is interesting, as experiments on the activity of amine oxidase in

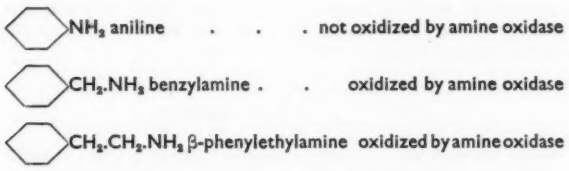
vitro might have led one to believe otherwise. For instance, in contrast to cell respiration, the oxidation of amines by amine oxidase in vitro is not inhibited by cyanide. However, there are some indications which suggest that, in the living organism, amine oxidase may not directly react with molecular oxygen, but that other hydrogen acceptors are interposed between the substrate and molecular oxygen (for a fuller discussion see Blaschko, 1952b). In the manometric experiment one finds that the enzymic activity depends upon the partial pressure of oxygen in the gas phase. Measurements are therefore usually carried out in an atmosphere of 100% oxygen; when one reduces the oxygen pressure to lower values such as occur in the living tissues, the rate of the enzymic reaction is considerably slowed down. This might be explained by assuming that, in the living cell, amine oxidase does not react with molecular oxygen, but that it does so in vitro. A similar behaviour of some flavin enzymes has been described.

The chemical composition of amine oxidase is unknown. The enzyme is not inhibited by substances that react with metals, e.g., by cyanide or carbon monoxide, but it is slightly inhibited by compounds that react with sulphydryl groups.

3. Substrate Specificity of Amine Oxidase

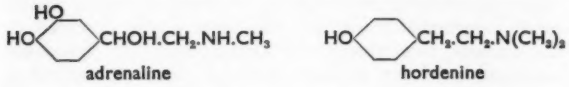
The substrate specificity of amine oxidase has been very extensively studied; this is because so many amines have marked biological activities and are of therapeutic interest.

Amine oxidase acts on amines which have an amino group attached to a terminal carbon atom, i.e., compounds of the type $—CH_2.NH_2$. Aromatic amines like aniline are not attacked, but benzylamine and β-phenylethylamine are oxidized:


NH₂ aniline . . . not oxidized by amine oxidase
CH₂.NH₂ benzylamine . . oxidized by amine oxidase
CH₂.CH₂.NH₂ β-phenylethylamine oxidized by amine oxidase

The free amino group is not essential for substrate specificity: methylamino compounds $—CH_2.NH.CH_3$ are attacked as rapidly as the corresponding primary amines. Adrenaline is an example of a secondary amine oxidized by amine oxidase.

The introduction of two *N*-methyl groups reduces the rate of the enzymic reaction considerably, but it does not altogether abolish it. Hordenine is a naturally occurring compound of this kind which is slowly oxidized by amine oxidase:



Quaternary ammonium bases are not oxidized by the enzyme.

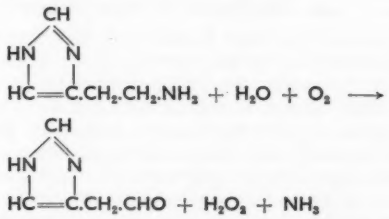
Aliphatic monoamines are oxidized by amine oxidase. In the homologous series: $CH_3(CH_2)_n.NH_2$ the rate of oxidation at first increases with n; with a further increase the rate falls again, probably because, with the long-chain monoamines, micelle formation decreases the effective substrate concentration in aqueous solution.

Short-chain aliphatic diamines, e.g., putrescine or cadaverine, are not oxidized by amine oxidase. The presence of two groups $—CH_2.NH_2$ in one and the same molecule interferes with the substrate-enzyme reaction:

A study of the enzymic oxidation of the homologous serie $H_2N.(CH_2)_n.NH_2$ has shown that this disturbing influence o the second basic group decreases as the distance between the two groups increases; the diamines with more than six methylene groups are oxidized by amine oxidase (Blaschko & Hawkins, 1950).

4. Histaminase and Amine Oxidase

In 1929 Best discovered an enzyme which brought about the biological inactivation of histamine. Histaminase probably catalyses an oxidative deamination reaction analogous to the amine oxidase reaction (see Mehler, Tabor & Bauer, 1952):



In 1938, Zeller showed that histaminase acted not only on histamine, but also on the short-chain diamines, putrescine and cadaverine (see Zeller, 1951).

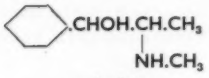
We can summarize the specificity requirements of the amine oxidases by saying that histaminase requires the presence in the substrate molecule of two basic groups; the optimal configuration is present when four or five atoms separate the two basic groups. On the other hand, the active centre of amine oxidase has only one site that reacts with a basic group. In addition, in order to account for the lack of affinity of the short-chain diamines for amine oxidase, we assume that amine oxidase requires an orientation of the substrate on the enzyme surface which the short-chain diamines cannot take up.

Although amine oxidase and histaminase catalyse reactions that are analogous, the two enzymes do not appear to be chemically closely related. Histaminase is inhibited by cyanide, but amine oxidase is cyanide-resistant. Histaminase is strongly inhibited by all carbonyl reagents, e.g., by semicarbazide, but amine oxidase is not inhibited by these reagents.

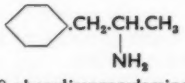
The fact that histaminase is inhibited by carbonyl reagents has led to the suggestion that this enzyme contains pyridoxal phosphate (Werle & Pechmann, 1949; Sinclair, 1952). There is also some evidence for the flavin nature of histaminase (Kapeller-Adler, 1949). Werle and Pechmann have put forward the idea that histaminase is a complex enzyme: (i) pyridoxal phosphate acts as an acceptor for the amino group; a transamination occurs and the aldehyde and pyridoxamine phosphate are formed; (ii) pyridoxamine phosphate is oxidized to pyridoxal phosphate by a flavoprotein. This picture will have to be reconciled with the observation of Zeller (1951) that histaminase oxidizes *NN'*-dimethylputrescine, a compound without a primary amino group. Pyridoxal enzymes usually react only with primary amino compounds.

5. Competitive Inhibitors of Amine Oxidase

Amines in which the amino group is not attached to a terminal carbon atom are not oxidized by amine oxidase, but many of these substances are competitive inhibitors of the enzyme. This was first described for ephedrine, but β-phenyl*iso*propylamine itself and many of its derivatives are competitive inhibitors:



ephedrine



β-phenyl*iso*propylamine (amphetamine, benzedrine)

Compounds without a side-chain hydroxyl group, like amphetamine, have a higher affinity for amine oxidase than compounds with a hydroxyl group, like ephedrine. In other words, amphetamine is a stronger inhibitor than ephedrine.

6. Distribution of Amine Oxidase

The enzyme has been found in all vertebrates examined. It occurs not only in the liver but in the whole alimentary tract and its glands. It is also present in the kidneys and in many other glands, endocrine and exocrine; it has been found in the uterus, in the placenta, in the male genital system and in the lungs. In the adrenal gland, enzymic activity is present in both cortex and medulla. The central nervous system contains amine oxidase and so do autonomic ganglia. The enzyme has also been found in the heart. In the blood vessels, amine oxidase occurs mainly in the arteries; less activity is found in the veins (Thompson & Tickner, 1951). More recently, Burn and Robinson (see Robinson, 1952) have described the occurrence of amine oxidase in the iris and in the muscle of the nictitating membrane in cats and rabbits.

7. Amine Oxidase and Amine Metabolism in the Living Organism

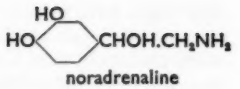
As already discussed, deamination by amine oxidase would account for the metabolic fate of tyramine in the animal body. The same is true for the many other amines which are substrates of the enzyme in vitro. Conversely, many amines that are not oxidized by amine oxidase are not broken down in the living organism. This is a fact which is of practical importance: ephedrine and ephedrine-like compounds owe their characteristic long-lasting action to their resistance to oxidation by amine oxidase; their action is terminated not by destruction but by the slower process of excretion.

8. Physiological Significance of Amine Oxidase

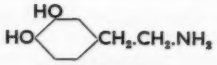
It is possible that in organs like the gut and the placenta amine oxidase may serve a "detoxicating" function, but there seems to be little doubt that the main task of the enzyme is in the metabolism of some of the substances for which Guggenheim (1951) has introduced the term "biogenic amines". These are the adrenaline-like compounds and the newly discovered tryptamine derivatives.

The three sympathicomimetic amines found in the body are adrenaline, noradrenaline and dopamine. (The term "dopamine" recently proposed by Dale is used here as it is less ambiguous than the term "hydroxytyramine".) It is now well established that the material found in adrenergic

adrenaline



noradrenaline



dopamine

nerves is noradrenaline (see von Euler, 1952). In many species, including man, the medullary hormone is mainly adrenaline, with some noradrenaline, and there is evidence that adrenaline is formed from noradrenaline in the adrenal medulla (Bülbring, 1949).

The possibility that the adrenergic transmitter is inactivated by amine oxidase has been discussed since Gaddum & Kwiatkowski (1938) studied the effect of ephedrine on the response of the blood vessels of the rabbit's ear not only to adrenaline, but also to sympathetic stimulation. New evidence in favour of this role of amine oxidase has since been obtained by Burn, who has recently reviewed it (Burn, 1952).

In manometric experiments, adrenaline and noradrenaline have a rather low affinity for amine oxidase, and this has raised doubts as to the physiological significance of amine oxidase in the biological inactivation of adrenaline and sympathin. However, it may be unwise to base estimates of the activity of the enzyme in vivo on the measurements in vitro. The affinity of other oxidases for their substrates increases as the partial pressure of oxygen decreases (Laser, 1952); it is likely that this would also be true for amine oxidase.

The isotopic tracer technique may eventually tell us more about the fate of the catechol amines in the body. Schayer, Smiley & Kaplan (1952) have established that some of the adrenaline molecules are split in the body between the β-carbon and the *N*-methyl carbon (the asterisks denote the carbon atoms labelled with ^{14}C):

These authors provisionally estimate that, of a small dose of adrenaline, about one-half is metabolized by amine oxidase.

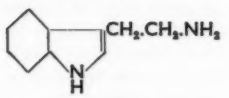
The third catechol amine found in the body is dopamine. It has been found in the urine, and increased amounts were excreted by a patient with phaeochromocytoma. Dopamine also occurs in the adrenal medulla of the sheep and the ox (Goodall, 1951; Shepherd & West, 1952). The enzyme L-DOPA decarboxylase, which forms dopamine from L-3 : 4-dihydroxyphenylalanine, has recently been found in the adrenal medulla of the ox (Langemann, 1951). Dopamine might therefore be considered as precursor of pressor amines in the suprarenal gland, but this view is not accepted by all authors (see von Euler, 1952), and the functional significance of the amine is not yet clear. Dopamine is rapidly oxidized by amine oxidase, and it has been pointed out by Polonovski and his colleagues (Polonovski, Schapira & Gonnard, 1946) that organs which contain DOPA decarboxylase are rich in amine oxidase.

9. Derivatives of Tryptamine

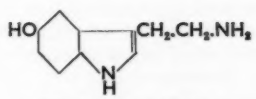
In 1934 Wieland, Konz & Mittasch established that bufotenine, a biologically active constituent of the skin secretions in toads, was the *N*-dimethyl derivative of 5-hydroxytrypt-

amine. It has since been established that 5-hydroxytryptamine itself occurs in mammalian tissues (Rapport, 1949).

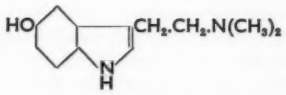
The amine is identical with serotonin, a substance isolated from cattle serum and believed to be responsible for some of its toxic actions. How it is formed is still unknown; a release from disintegrating platelets during clotting has been discussed.



tryptamine



5-hydroxytryptamine
(serotonin, enteramine)



bufotenine

For over ten years, Erspamer has studied the properties of enteramine, an active material present in extracts of the intestinal mucosa. Enteramine is now known also to be identical with 5-hydroxytryptamine (Erspamer & Asero, 1951). Erspamer considers the amine as a true tissue hormone, the product of the enterochromaffin cells of the gastro-intestinal tract. Recent work at the National Institute for Medical Research at Mill Hill, London, has shown that intestinal extracts contain another substance related to 5-hydroxytryptamine (Dalgliesh, Toh & Work, 1952).

The biological significance of this new " biogenic " amine remains to be established. Like tryptamine, it is rapidly oxidized by amine oxidase (Freyburger *et al.* 1952; Blaschko, 1952a). The high rate of oxidation suggests that the amine is also an important substrate of amine oxidase in vivo. In the oxidation of 5-hydroxytryptamine, coloured oxidation products are formed. Similar products had already been seen in the biological oxidation of tryptamine (Pugh & Quastel, 1937b); the colour probably arises in secondary reactions of the aldehyde formed as the primary oxidation product.

10. Amine Oxidase in Invertebrates

True chromaffin tissue occurs in all vertebrates, but chromaffin cells have also been found in annelid worms. In the earthworm, the inhibitory nerve supply to the gut appears to be adrenergic (Millott, 1943). We have recently found that homogenates of the gut of the earthworm have amine oxidase activity. Homogenates of the worm that do not include gut have very little activity. Here it is the adrenergically innervated structure that contains the enzyme.

Another group in which amine oxidase is associated with amines are the cephalopods. The livers of octopus and sepia are the richest sources of enzyme yet found (Blaschko & Hawkins, 1952); the enzyme also occurs in many other organs. The octopods contain not only 5-hydroxytryptamine, but also tyramine and related compounds. It seems therefore that the distribution of amine oxidase in nature is such that the enzyme is present wherever amines are found upon which it is capable of acting.

REFERENCES

Best, C. H. (1929) *J. Physiol.* **67,** 256
Blaschko, H. (1952a) *Biochem. J.* **52,** *Proc.* x
Blaschko, H. (1952b) *Pharmacol. Rev.* **4,** 415
Blaschko, H. & Hawkins, J. (1950) *Brit. J. Pharmacol.* **5,** 625
Blaschko, H. & Hawkins, J. (1952) *J. Physiol.* **118,** 88
Blaschko, H., Richter, D. & Schlossmann, H. (1937a) *J. Physiol.* **90,** 1
Blaschko, H., Richter, D. & Schlossmann, H. (1937b) *Biochem. J.* **31,** 2187
Bülbring, E. (1949) *Brit. J. Pharmacol.* **4,** 234
Burn, J. H. (1952) *Brit. med. J.* **1,** 784
Cotzias, G. C. & Dole, V. P. (1951) *Proc. Soc. exp. Biol., N.Y.* **78,** 157
Dalgliesh, C. E., Toh, C. C. & Work, T. S. (1952) *Biochem. J.* **52,** *Proc.* xxx
Erspamer, V. & Asero, B. (1951) *Ric. sci.* **21,** 2132
Euler, U. S. von (1952) *II^e Congrès International de Biochimie, Paris, 21–27 Juillet, 1952.* In: *Symposium sur les hormones protéiques et dérivées des protéines.* Société d'Edition d'Enseignement Supérieur, Paris
Ewins, A. J. & Laidlaw, P. P. (1910) *J. Physiol.* **41,** 78
Freyburger, W. A., Graham, B. E., Rapport, M. M., Seay, P. H., Govier, W. M., Swoap, O. F. & Vander Brook, M. J. (1952) *J. Pharmacol.* **105,** 80
Gaddum, J. H. & Kwiatkowski, H. (1938) *J. Physiol.* **94,** 87
Goodall, M. (1951) *Acta physiol. scand.* **24,** Suppl. No. 85
Guggenheim, M. (1951) *Die biogenen Amine, und ihre Bedeutung für die Physiologie und Pathologie des pflanzlischen und tierischen Stoffwechsels,* 4th ed. Karger, Basle & New York
Hare, M. L. C. (1928) *Biochem. J.* **22,** 968
Hawkins, J. (1952) *Biochem. J.* **50,** 577
Kapeller-Adler, R. (1949) *Biochem. J.* **44,** 70
Kohn, H. I. (1937) *Biochem. J.* **31,** 1693
Langemann, H. (1951) *Brit. J. Pharmacol.* **6,** 318
Laser, H. (1952) *Proc. roy. Soc.* B, **140,** 230
Mehler, A. H., Tabor, H. & Bauer, H. (1952) *J. biol. Chem.* **197,** 475
Millott, N. (1943) *Proc. roy. Soc.* B, **131,** 362
Polonovski, M., Schapira, G. & Gonnard, P. (1946) *Bull. Soc. Chim. biol., Paris,* **28,** 735
Pugh, C. E. M. & Quastel, J. H. (1937a) *Biochem. J.* **31,** 286
Pugh, C. E. M. & Quastel, J. H. (1937b) *Biochem. J.* **31,** 2306
Rapport, M. M. (1949) *J. biol. Chem.* **180,** 961
Robinson, J. (1952) *Brit. J. Pharmacol.* **7,** 99
Schayer, R. W., Smiley, R. L. & Kaplan, E. H. (1952) *J. biol. Chem.* **198,** 545
Shepherd, D. M. & West, G. B. (1952) *J. Physiol.* **117,** 67*P*
Sinclair, H. M. (1952) *Biochem. J.* **51,** *Proc.* x
Thompson, R. H. S. & Tickner, A. (1951) *J. Physiol.* **115,** 34
Werle, E. & Pechmann, E. von (1949) *Liebigs Ann.* **562,** 44
Werle, E. & Roewer, F. (1952) *Biochem. Z.* **322,** 320
Wieland, H., Konz, W. & Mittasch, H. (1934) *Liebigs Ann.* **513,** 1
Zeller, E. A. (1951) In: Sumner, J. B. & Myrbäck, K. D. R., ed. *The enzymes : chemistry and mechanism of action,* vol. 2, pt 1, p. 536. Academic Press, New York

ADAPTIVE ENZYMES

SIR CYRIL HINSHELWOOD D.Sc. F.R.S.

Physical Chemistry Laboratory
University of Oxford

1 Adaptation and growth
2 Mutations and adaptive changes
3 Reversibility of adaptation
4 Genetical theory in relation to adaptation
5 Direct adaptive response
References

Enzymes are catalytic agents built up in cells in the general course of cell growth, and containing protein as at any rate a major constituent. Their mode of action must be rather complex and frequently they require the presence not only of the substance whose reaction they catalyse, but also of a diffusible metabolite or coenzyme. The enzymic properties of some cell preparations are manifested almost immediately the cells are brought into contact with their substrate, but in other cases there may be a long period of induction before the required action comes into play. On the strength of this fact there was introduced the idea of a distinction between constitutive and adaptive enzymes, the former being essential to the make-up of a cell, the latter called into being by the substrate upon which they act (Karström, 1938; Dubos, 1940).

Yudkin (1938) interpreted this phenomenon in terms of an equilibrium between enzymes and precursors, which would be displaced in the required direction by the presence of the substrate combining with the enzyme. This explanation may well be correct in some cases, but a more general one seems necessary to do justice to the whole set of facts now known. In particular, it is clear that the gradual development of an enzymic activity is simply one manifestation of the time required by many complex chemical reactions to reach what is known as their steady state. Chemical reactions of the most varied kinds which take place in a series of stages reach their maximum velocity after a finite time—the induction period—and there is no reason why a complex enzyme reaction or set of enzyme reactions should not follow this general law.

Numerous examples of the adaptive systems have been studied and only a few typical examples can be cited. Stephenson & Gale (1937) studied galactose fermentation by washed suspensions of *Bacterium coli*, and Stephenson & Yudkin (1936) galactose fermentation by yeast cultures. Hegarty (1939) examined the development of the power of *Streptococcus lactis* to attack various sugars. The reduction of tetrathionate (Pollock & Knox, 1943; Knox & Pollock, 1944; Jebb, Knox & Tomlinson, 1950), and also the reduction of nitrate (Pollock & Wainwright, 1948), have been the subject of much work.

1. Adaptation and Growth

Some authors have stated that the development of the enzyme is conditioned by the actual multiplication of the cells in presence of the substrate, and other authors (or sometimes the same authors working with a different system) believe that the induction of the enzymic activity anticipates the growth of the cells.

They may here have been a little at cross-purposes in terminology. Clearly, if the enzyme reaction involves a fairly complex sequence in which the intermediates require time to be built up, and if growth does not ensue until this enzyme activity has become established to at least a minimum level, then what might be called adaptation precedes growth. If, on the other hand, a period of active growth so establishes the amount or activity of the enzyme that the long lag is not necessary on a subsequent occasion, then we may well speak of a growth-linked adaptation. There is really no sharp distinction between the two cases, but it might be well to call an initial period of delay an induction period, and to use the term adaptation for the establishment during one period of growth of a state of affairs such that subsequent induction periods are shorter than before or entirely absent.

Pollock & Wainwright (1948) found that the capacity of cells to reduce nitrate developed before the start of growth. On the other hand, they found that specific treatment with tetrathionate or nitrate conferred a corresponding advantage in respect of subsequent growth-rate in media containing tetrathionate or nitrate respectively.

There are various possible explanations (not mutually exclusive) about what may be happening during the induction period and during the subsequent phase of adaptive growth. The simplest kind of adjustment would consist simply in the building up of the proper kinds of intermediates in the cells in the right proportions, in the way which is familiar in the chemical kinetics of, for example, chain reactions. Changes in the concentration of such intermediates certainly seem to play an important part in some cell phenomena. For example, when cultures of *Bact. lactis aerogenes* in active aerobic growth are made anaerobic there is a sudden cessation of growth. During the arrest the reducing power increases but falls again approximately to its original value as anaerobic growth gradually sets in. On the other hand, re-aeration of an arrested anaerobic culture leads to almost immediate resumption of optimum growth. Baskett & Hinshelwood (1951a) interpreted such observations in the light of the view that transition from aerobic to anaerobic conditions involved the development of a new mechanism for the re-oxidation of reduced hydrogen carriers, but that re-aeration simply lowered the concentrations of certain of these.

Some sort of process by which in adaptive phenomena the actual molecular texture of the enzyme proteins is changed has sometimes been suggested (e.g., Spiegelman & Dunn, 1947), and Cohn (quoted by Monod, 1951) has shown that the adaptation of *Bacterium coli* to lactose is accompanied by the appearance of one specific antigen and the partial disappearance of a similar but different one.

Changes in enzyme content of cells grown under different conditions of pH were found by Gale & Epps (1942)[1], the indication being that in some cases the amount of enzyme increased as the pH moved away from the optimum for activity. On the other hand, Eddy (unpublished results) has found that during the death of cultures in highly acid conditions a certain protective adjustment can occur, and this must be conditioned by internal metabolic activity not linked

[1] See also Gale, p. 135 of this Bulletin.

with growth. Virtanen (1949) has shown that the enzyme activity of cells changes as their nitrogen content is altered by cultivation under varying conditions of nitrogen supply, so that here quantitative adjustments are clearly indicated.

2. Mutations and Adaptive Changes

Adaptive phenomena are most clearly and revealingly shown when bacterial cultures (or cultures of other unicellular organisms) are subjected to serial subculture in fresh media containing either new sources of carbon or nitrogen or inhibitory drugs. On successive subculture long lags initially shown disappear and growth-rates progressively improve to optimum values. Various specific functions concomitantly improve to optima. On the one hand, it is natural to consider these phenomena on the same footing as the study of induction phenomena in non-growing enzyme suspensions: on the other hand, adaptive interpretations come, at first sight, into head-on collision with the views of geneticists on mutation and selection. There is, however, no necessary conflict.

The mutation theory is simply stated. All cell properties depend upon genes which are copied in inheritance, and change in character only by discontinuous mutations. These occur by chance in a completely random fashion and are in no way conditioned by the action of the agent to which the cells appear to become adapted. The adaptation is merely the selection of the appropriate mutants. The postulated mutations are multitudinous and show gradations indistinguishable from a continuous variation. They must therefore involve not only very many genes in all, but many for each single kind of property (polygenic systems). Mutations and reverse mutations have to be assumed to be in a lively dynamic equilibrium.

The adaptive theory takes the view that the properties which develop do so in response to an action of the specific substrate or drug, and here there are several possibilities. The enzymic make-up of the cell may become changed, either qualitatively, a new kind of material (for example, a somewhat differently folded protein) appearing; or quantitatively, the proportions of existing enzymes being more or less profoundly altered. The antigen-antibody relation shows that induced changes of the first kind are not at all unlikely in principle, and induced changes of the second kind can be shown not only to be possible according to the principles of reaction kinetics but, on the basis of some general and plausible assumptions about cell growth, to be necessary and automatic (Hinshelwood & Lewis, 1948; Hinshelwood, 1952).

It may be observed here that, were some small fraction of the total enzyme content of a cell to be more actively catalytic than the rest towards a given substrate, the same kind of principles show that this fraction would expand relatively to the rest in the autosynthetic growth of the cell. Thus it can be seen that this " internal selection " of some part of the cell substance is in one way not so very different from the " external selection " of some mutant cells in a population.

Mutations and adaptive changes in cell organization are not incompatible. Cells of micro-organisms, even when they have not clearly defined nuclei, have in all probability physically and chemically differentiated regions, the structure of which stands in an organic relation with the make-up of the cell as a whole. Not only does the whole cell divide, but any mechanically separable parts of it divide also, and in this process one can imagine some statistical variation in the distribution of material, in the precise folding of proteins and even in the way chemical linkages are arranged. A sufficiently unusual chance might give rise to a form which would continue to be copied in subsequent autosynthesis, and which might rank as a mutant. Irradiation with ultraviolet light and treatment with various chemicals certainly do give rise to a proportion of cells with changed properties, but in practically every known case, if not in all, the change is the loss of a property rather than the gain of a new one. This means that the chemical picture of a mutation leading to an adaptively useful property is rather difficult to envisage, and from the physico-chemical point of view an adaptation arising from interaction with the substrate is an easier one.

It is also in some respects a more natural one, because it is predictable. Cells grow by the functioning of an elaborate network of chemical reactions, and there is much evidence that numerous alternatives exist for the exact pattern followed. On the basis of quite simple and general principles it can be shown that the mode of reaction leading to the optimum growth-rate in given circumstances should automatically be established. To take a simple example: if an antibiotic agent interferes with the use of a certain metabolite, there should occur an adjustment of the enzymic proportions which will yield an increased supply of this metabolite to compete with the toxic substance (Pryce, Davies & Hinshelwood, 1944, 1945).

3. Reversibility of Adaptation

The physico-chemical view predicts that adaptive changes should be reversible, but that in certain circumstances the reversion may occur so slowly that a measure of permanence is simulated. The experimental facts are that sometimes the changes are quite easily reversible and sometimes very nearly stable. (By reversibility is meant this: if cells are adapted to resist an antibiotic and are then subcultured in its absence, they lose their resistance. Often, however, they may be cultivated for a very long time away from the antibiotic but still show resistance when finally tested again in its presence.)

The usual sequence of events is as follows: after cultivation in the new conditions just long enough to attain a nearly optimum growth-rate, return to the original medium leads to reversal, but the adaptation becomes more stable the longer it is impressed.

This might be thought to be explained by the progressive purification of a mutant strain from an original form which regains its preponderance on culture in the original medium, unless it has been completely eliminated from the mixture. But experiments with artificially mixed populations of adapted and unadapted cells show that this cannot be the general explanation (Baskett & Hinshelwood, 1951b).

The adaptation to one substrate can usually be impaired and at least partially driven out by subsequent adaptation to a series of others (Scopes & Hinshelwood, 1952), but in some cases multiple adaptations—for example, to more than one drug or to a series of related sugars—can be held simultaneously (Postgate & Hinshelwood, 1946).

4. Genetical Theory in Relation to Adaptation

The frequently high degree of persistence of drug resistance, or of substrate adaptation—for example, of *Bact. coli mutabile*

to lactose (Massini, 1907)—has been used by the supporters of the mutation view to dismiss the idea of adaptation as " Lamarckian ", on the ground that it represents the " inheritance of an acquired character ". Apart from the question whether appeals to authoritative dogma and the citation of names constitutes scientific argument, this objection is illogical for several reasons. First, what constitutes " inheritance " in cells which multiply by simple division is not the same as in sexually reproducing organisms, and the persistence of the character, although often great, is not absolute. Secondly, a change in chemical linkages, protein folding or the like, produced in a bacterium by chance, is just as much " acquired " as one caused by a known process of adjustment. The issue of mutations and automatic adjustments must be settled by experiment. Both are possible in principle, and both may occur in actuality. On the whole, one might expect the commoner changes of limited amplitude to be due to adaptive adjustments, rarer changes of a more profound kind to have to await exceptional mutations. Certainly nothing comparable to a species change which might have been attributable to either agency has been experimentally observed in bacteria.

There is, it should be observed, no conflict between Mendelian principles and theories of adaptive changes. In the sporulation and mating of yeast cells, clear-cut Mendelian phenomena occur (Winge & Roberts, 1948, 1950; Lindegren, 1949), but superposed on them occur very well-defined adaptive changes which are almost certainly not due to mutations (Kilkenny & Hinshelwood, 1951, 1952), but simply to changed enzymic organization of the kind already referred to.

Experimental studies have been made by various methods, with the object of deciding in individual cases whether mutations or induced changes underlie apparent adaptations. The well-known fluctuation test of Luria & Delbrück (1943) has been widely applied. The principle is this: the statistical variation of, for example, drug resistance among a number of similar cultures derived from the same small inoculum, is compared with that among simultaneous samples from one given culture. If mutations have been occurring in a random fashion, the first set should show a greater variance. It nearly always does. The method has been criticized (Eriksen, 1949; Dean & Hinshelwood, 1952) on the ground that it neglects all other causes of biological differences between the separate cultures, and that such causes do in fact exist. Indeed it is doubtful whether, if this method had not yielded an answer in accordance with a certain preconceived idea, it would have been so widely accepted without criticism in the first place.

The method of " replica plating " (Lederberg & Lederberg, 1952) has given interesting results with streptomycin resistance. In this a resistant colony can be recognized not by its behaviour in presence of the drug, but by the behaviour of a corresponding colony on a parallel plate. (Replicate inoculations to several plates are made by means of a stamping device covered with a sterile fabric.) In this way suspected resistant types can be selected and subcultured without contact with the drug. Lederberg and Lederberg state that they have obtained in this way types highly resistant to streptomycin, and cite the evidence as confirmation of " participation of spontaneous mutation " in the adaptation.

In connexion with what was previously said about the rarity and the amplitude of mutations it is of interest to note that in this case the original streptomycin-resistant type was so exceptional that with inocula of over 10^9 cells only two or three mutants occurred with 20–30 plates, and that only one of these was enriched by selection in the way described. This leaves open the question of what would happen with degrees of streptomycin resistance lower and more usual than that studied here. In the case of other drugs and more moderate concentration ranges there is good evidence for direct adaptive response. The resistance is precisely graded according to the concentration of drug to which the cells have been " trained ", and the stability of the resistance is variable. Moreover, even for a given drug concentration the resistance develops in a continuous and gradual manner. Such facts would demand elaborate hypotheses about multiple mutations of different kinds, which are not at all probable. Strains resistant to streptomycin and capable of growth in a medium without specific amino-acids which they originally demand can be produced by training them first to dispense with the amino-acid and then to resist the drug (Sevag & Rosanoff, 1952). Pre-existent mutants with these properties could be selected in one operation. Here again the mutation hypothesis would demand assumptions of a purely ad hoc type.

5. Direct Adaptive Response

There is good evidence for the direct adaptive response of coliform bacteria to carbohydrates and other nutrient sources which they are at first able to assimilate only after a long lag phase. The example of *Bact. lactis aerogenes* and D-arabinose has been studied in detail (Koser & Vaughan, 1937; Jackson & Hinshelwood, 1947, 1949, 1950; Baskett & Hinshelwood, 1951b). Another good example is the ability of various yeast strains to utilize galactose (Winge & Roberts, 1948, 1950; Kilkenny & Hinshelwood, 1951, 1952). The cells show very long lags at a first transfer to a medium containing the new compound as sole carbon source. If this lag represented the time required for preformed mutants to multiply, they would be present in a proportion of somewhere about one in 10^6, or less. But the cells present show a total increase in mass of perhaps 50% before any detectable increase in number occurs. Therefore in such cases the bulk of the cell population is showing a response, and it is presumably the adaptive response since it is immediately followed by growth.

Growth will occur from inocula so small that the probability of their containing mutants is negligible: preliminary incubation in a liquid medium containing D-arabinose shortens the time required for the subsequent development of individual colonies of *Bact. lactis aerogenes* when they are grown on a solid agar plate in presence of D-arabinose. It is not merely the proportion of positive results that is increased in this way. The first cultures (of the bacteria or yeasts) do not yield fully trained strains as a simple selection would do, but only partially trained ones: reversion does not follow the quantitative pattern which a theory of back-mutation would suggest, and the development, or loss, of power to utilize the carbon source may occur in too few generations to provide sufficient time for selection.

In view of these and other facts it would seem that the adaptive changes of enzymic organization, which should in principle be possible, do in fact occur. Whether, in parallel cases, discontinuous mutations also play their part, and if so, how frequently, by what mechanism and to what extent, are problems still largely lacking agreed solutions.

REFERENCES

Baskett, A. C. & Hinshelwood, C. N. (1951a) *Proc. roy. Soc.* B, **138,** 75
Baskett, A. C. & Hinshelwood, C. N. (1951b) *Proc. roy. Soc.* B, **139,** 58
Dean, A. C. R. & Hinshelwood, C. N. (1952) *Proc. roy. Soc.* B, **139,** 236
Dubos, R. (1940) *Bact. Rev.* **4,** 1
Eriksen, K. R. (1949) *Acta path. microbiol. scand.* **26,** 269
Gale, E. F. & Epps, H. M. R. (1942) *Biochem. J.* **36,** 600
Hegarty, C. P. (1939) *J. Bact.* **37,** 145
Hinshelwood, C. N. (1952) *J. chem. Soc.*, p. 745
Hinshelwood, C. N. & Lewis, P. R. (1948) *Proc. roy. Soc.* B, **135,** 316
Jackson, S. & Hinshelwood, C. N. (1947) *Trans. Faraday Soc.* **44,** 568
Jackson, S. & Hinshelwood, C. N. (1949) *Proc. roy. Soc.* B, **136,** 562
Jackson, S. & Hinshelwood, C. N. (1950) *Proc. roy. Soc.* B, **137,** 88
Jebb, W. H. H., Knox, R. & Tomlinson, A. H. (1950) *J. gen. Microbiol.* **4,** 374
Karström, H. (1938) *Ergebn. Enzymforsch.* **7,** 350
Kilkenny, B. C. & Hinshelwood, C. N. (1951) *Proc. roy. Soc.* B, **139,** 73
Kilkenny, B. C. & Hinshelwood, C. N. (1952) *Proc. roy. Soc.* B, **139,** 575
Knox, R. & Pollock, M. R. (1944) *Biochem. J.* **38,** 299
Koser, S. A. & Vaughan, E. F. (1937) *J. Bact.* **33,** 587
Lederberg, J. & Lederberg, E. M. (1952) *J. Bact.* **63,** 399
Lindegren, C. C. (1949) *The yeast cell.* Educational Publishers, St Louis, Missouri
Luria, S. E. & Delbrück, M. (1943) *Genetics,* **28,** 491
Massini, R. (1907) *Arch. Hyg., Berl.* **61,** 250
Monod, J. (1951) *Bull. World Hlth Org.* **6,** 75
Pollock, M. R. & Knox, R. (1943) *Biochem. J.* **37,** 476
Pollock, M. R. & Wainwright, S. D. (1948) *Brit. J. exp. Path.* **29,** 223
Postgate, J. R. & Hinshelwood, C. N. (1946) *Trans. Faraday Soc.* **42,** 45
Pryce, J. M. G., Davies, D. S. & Hinshelwood, C. N. (1944) *Trans. Faraday Soc.* **40,** 397
Pryce, J. M. G., Davies, D. S. & Hinshelwood, C. N. (1945) *Trans. Faraday Soc.* **41,** 163; 465; 778
Scopes, A. W. & Hinshelwood, C. N. (1952) *J. chem. Soc.*, p. 1838
Sevag, M. G. & Rosanoff, E. (1952) *J. Bact.* **63,** 243
Spiegelman, S. & Dunn, R. (1947) *J. gen. Physiol.* **31,** 153
Stephenson, M. & Gale, E. F. (1937) *Biochem. J.* **31,** 1311
Stephenson, M. & Yudkin, J. (1936) *Biochem. J.* **30,** 506
Virtanen, A. I. (1949) *Ann. Acad. Sci. fenn.* (Series AII, Chemistry) No. 36
Winge, O. & Roberts, C. (1948) *C. R. Lab. Carlsberg* (Sér. Physiol.) **24,** 277
Winge, O. & Roberts, C. (1950) *C. R. Lab. Carlsberg* (Sér. Physiol.) **25,** 36
Yudkin, J. (1938) *Biol. Rev.* **13,** 93

* * *

ENZYMES ACTING ON NUCLEIC ACIDS

J. N. DAVIDSON M.D. D.Sc.

Gardiner Professor of Physiological Chemistry
University of Glasgow

1 Ribonucleic acids, nucleotides and nucleosides
2 Ribonuclease
3 Structure of ribonucleic acid (RNA)
  a Evidence from chemical behaviour
  b Evidence from enzymic breakdown
  c The occurrence of the $C_{(5')}$-linkage
  d The occurrence of branched chains
4 Deoxyribonucleic acid (DNA) and deoxyribonuclease
5 Clinical uses of deoxyribonuclease
References

It is now well known that two main groups of nucleic acids are to be found in the living cell, the ribonucleic acid (RNA) group and the deoxyribonucleic acid (DNA) group. The latter occurs almost exclusively in the cell nucleus, where in combination with nuclear proteins it forms the deoxyribonucleoprotein known to the histologists as chromatin. In the nucleus the ribonucleic acids are found only in small amounts, chiefly in the nucleolus, but they are abundant in the cell cytoplasm where, as ribonucleoproteins, they form the chief basophilic constituents of the cytoplasmic granules (the large granules or mitochondria and the small granules or microsomes) and of the non-sedimentable cytoplasmic substance or cell sap. The biochemistry of these nucleic acids has recently been extensively reviewed (Davidson, 1950, 1953; Baddiley, 1951a, 1951b; Jordan, 1952), but some of the salient features will be emphasized here.

1. Ribonucleic Acids, Nucleotides and Nucleosides

Ribonucleic acid (RNA) is most conveniently prepared from yeast, but it can readily be obtained also from mammalian tissues, especially liver, and from plant viruses such as tobacco mosaic virus. Since RNA is difficult to isolate without some degradation of the macromolecules and since it generally forms polydisperse solutions on which sedimentation constants and diffusion constants are difficult to determine, figures for the molecular weights of different ribonucleic acids must be regarded as approximations, but they are known to be very high and values as great as 300,000 have been recorded.

When RNA is broken down by hydrolysis the products shown in Table I are obtained. Treatment with dilute alkali yields partial breakdown products known as *nucleotides*, each of which consists of the pentose sugar ribose condensed with a purine or pyrimidine base through a glycosidic linkage at carbon atom number 1, $C_{(1')}$, on the ribofuranose ring and esterified with phosphoric acid at carbon atom number 3, $C_{(3')}$, or at number 2, $C_{(2')}$ (Cohn, 1951). One such nucleotide (adenylic acid), in which the base is adenine, is shown in fig. 1. Removal of the phosphoric acid grouping yields a *nucleoside*, in this case adenosine. The adenylic acid shown in fig. 1 is adenosine-3′-phosphate. Adenosine-5′-phosphate is also known to occur naturally and is sometimes referred to as muscle adenylic acid.

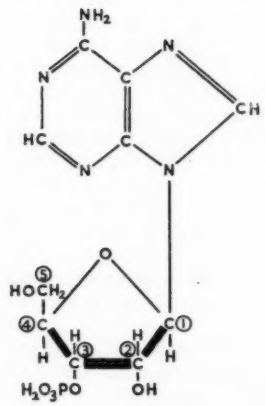
TABLE I. HYDROLYSIS PRODUCTS OF THE NUCLEIC ACIDS

Component	RNA	DNA
Purine bases	Adenine Guanine	Adenine Guanine
Pyrimidine bases	Cytosine Uracil	Cytosine Thymine (5-methyl uracil) 5-Methyl cytosine
Sugar	D-Ribose	D-2-Deoxyribose
Inorganic acid	Phosphoric acid	Phosphoric acid

The nucleotides formed from the other bases in RNA are known as guanylic acid, cytidylic acid and uridylic acid, while their corresponding nucleosides are guanosine, cytidine and uridine. The synthesis of the nucleotides and nucleosides by Professor A. R. Todd and his colleagues in Cambridge is one of the major recent achievements of British organic chemistry (Baddiley, 1951a, 1951b).

It used to be thought that RNA contained equimolecular proportions of the four purine and pyrimidine bases and that the structure was built up on the basis of a tetranucleotide. This tetranucleotide hypothesis has now had to be abandoned since recent methods of analysis have revealed that the molar proportions of bases in specimens of RNA from different biological sources show a very wide range of variation (Table II). These methods depend upon the separation of the hydrolysis products of RNA by the technique of paper chromatography as adapted for purine and pyrimidine bases and their nucleotides by Chargaff (1950) and his colleagues

FIG. 1. ADENOSINE-3′-PHOSPHATE, ONE FORM OF ADENYLIC ACID



The numbers in circles denote the 1′- to 5′-positions in the ribose molecule.

TABLE II. MOLAR PROPORTIONS, RELATED TO ADENINE AS 10, IN RIBONUCLEIC ACID (RNA) AND DEOXYRIBONUCLEIC ACID (DNA) FROM VARIOUS SOURCES AS DETERMINED BY VARIOUS AUTHORS

Source	Adenine	Guanine	Cytosine	Uracil	Thymine	5-Methyl cytosine
RNA						
Yeast	10	12.1	7.8	9.0	—	—
Tobacco mosaic virus	10	9.0	6.1	9.1	—	—
Rat liver	10	16.9	14.7	10.8	—	—
DNA						
Calf thymus	10	7.6	7.5	—	9.8	0.46
Wheat germ	10	8.9	6.5	—	10.2	2.2
Tubercle bacillus	10	15.8	18.6	—	11.0	0.0
Bacteriophage T_5	10	6.1	3.1	—	10.7	0.0

in New York and by Markham & Smith (1949, 1950, 1951a, 1952) in Cambridge.

It is clear from Table II that there must be many different forms of RNA but there is no reason to believe that they are not all built up in the same way by the condensation of large numbers of nucleotides to form a polynucleotide chain, probably with frequent branches. The nature of the internucleotide linkages involved in the formation of the polynucleotides has been worked out chiefly for yeast RNA, which is most readily obtainable in bulk, and largely with the aid of enzymes, of which the most important is ribonuclease.

2. Ribonuclease

Ribonuclease (RNase) is the best known of the nucleodepolymerase group of enzymes which hydrolyse nucleic acids without liberating free phosphate. It was first described by Jones in 1920 and was prepared by Kunitz (1940) in crystalline form from acid extracts of ox pancreas. The crystalline enzyme as usually prepared tends to be slightly contaminated with proteolytic impurities which are completely eliminated by the improved method of preparation described by McDonald (1948). Only such protease-free enzyme preparations should be used for critical work and especially for histochemical tests.

Crystalline ribonuclease is a protein of the albumin type with a molecular weight of about 15,000. It is remarkably resistant to heat in slightly acid solution although it is readily inactivated by alkali.

The exact mode of action of ribonuclease is still far from clear. Digestion of yeast RNA with the enzyme is accompanied by the gradual formation of free acidic groups without the liberation of free phosphoric acid. When the action is allowed to go to completion a non-dialysable residue or "core" of the nucleic acid molecule in the form of a "limit polynucleotide" is produced. It is resistant to the further action of ribonuclease and contains a large proportion of the total purines of the RNA. Its exact nature is discussed in the next section.

Ribonuclease has recently been made use of in two major applications. The first of these is the histochemical test first described by Brachet (1940), who showed that, when a tissue section is incubated with an aqueous solution of ribonuclease, RNA is removed so that those parts of the cell in which it was originally located cease to stain with basic dyes such as pyronine or toluidine blue, or cease to absorb ultra-violet light as demonstrated by photomicrography with the quartz microscope (Davidson & Waymouth, 1946).

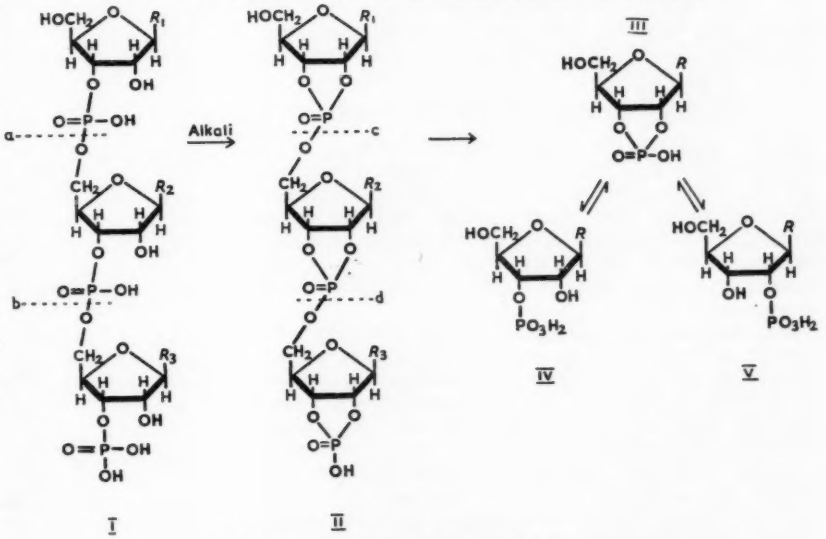
The second application of ribonuclease is in the elucidation of important structural features of RNA by Todd and his colleagues and by Markham and Smith in Britain, and by Chargaff and by Cohn in the USA.

3. Structure of Ribonucleic Acid (RNA)

a. Evidence from Chemical Behaviour

It has already been mentioned that alkaline hydrolysis of RNA yields the two isomeric nucleotides, adenosine-2′-phosphate and adenosine-3′-phosphate, which are also known as adenylic acids *a* and *b*, although not necessarily respectively. They may be represented by formulae (IV) and (V) in fig. 2. The isomers are readily converted into a mixture of both under acid conditions but are stable without interconversion in alkaline solution. Todd and his collaborators have shown that interconversion involves the formation of a cyclic intermediate, the nucleoside-2′ : 3′-phosphate (III, fig. 2), which yields on hydrolysis a mixture of the 2′- and 3′-phosphates (Brown & Todd, 1952). Such cyclic phosphates have been

FIG. 2. HYDROLYSIS OF A TRINUCLEOTIDE BY ALKALI



R represents a purine or pyrimidine base.

synthesized by Brown, Magrath & Todd (1952). They are also found during the alkaline hydrolysis of ribopolynucleotides. For example, the trinucleotide (I) shown in fig. 2, owing to the proximity of a hydroxyl group to a phosphoryl group, yields the cyclic intermediate (II) which is further degraded by the ready fission of the triester groups exclusively at the P-O-$C_{(5')}$ linkages (at c and d) with the formation of cyclic nucleoside-2′: 3′-phosphates (III) which then give rise to a mixture of nucleoside-2′- and 3′-phosphates (IV and V).

b. Evidence from Enzymic Breakdown

The action of ribonuclease also provides evidence on RNA structure. The products of enzymic hydrolysis consist of (i) pyrimidine mononucleotides representing a large proportion of the pyrimidine content of the RNA, (ii) easily diffusible di- and tri-nucleotides, (iii) a ribonuclease-resistant nucleic acid residue or "limit polynucleotide" forming what has been termed the "core" of the RNA molecule (Magasanik & Chargaff, 1951). This core probably consists of a mixture of tetra- and penta-nucleotides, each composed of purine nucleotides with a terminal pyrimidine nucleotide (Markham & Smith, 1952). It readily dialyses through cellulose membranes against salt solutions but not against water.

It is clear that the main action of ribonuclease on RNA is the splitting off of pyrimidine nucleotides. Some of these probably arise from side chains but others may be produced from the main chain. Thus the trinucleotide (I, fig. 2) will be split at point a if R_1 is a pyrimidine residue and at b if R_2 is a pyrimidine residue. The action of ribonuclease is to split the linkage joining the phosphate residue at $C_{(3')}$ in a pyrimidine nucleotide to $C_{(5')}$ in the next nucleotide in sequence, and Markham & Smith (1952) regard ribonuclease as a highly specific phosphodiesterase which will hydrolyse only secondary phosphate esters of pyrimidine nucleoside-3′-phosphates. It will, therefore, also hydrolyse the cyclic 2′: 3′-secondary phosphates of the pyrimidine nucleosides, as will be seen later (fig. 4).

To take a slightly more elaborate case, the polynucleotide PuPyPyPyPyPuPuPuPuPyPyPyPuPuPyPy (where Pu represents a purine nucleotide residue and Py a pyrimidine nucleotide residue) would be split by ribonuclease as shown by the vertical strokes thus:

PuPy|Py|Py|Py|PuPuPuPuPy|Py|Py|PuPuPy|Py

to yield: PuPy + PuPuPuPuPy + PuPuPy + 6 Py,
i.e., 6 pyrimidine nucleotides, viz., a dinucleotide, a trinucleotide, and a pentanucleotide, each with a terminal pyrimidine nucleotide residue linked to $C_{(3')}$ of the preceding purine nucleotide but composed otherwise of purine nucleotide residues. Such di- and tri-nucleotides have in fact been isolated by Markham & Smith (1952) from the products of ribonuclease action.

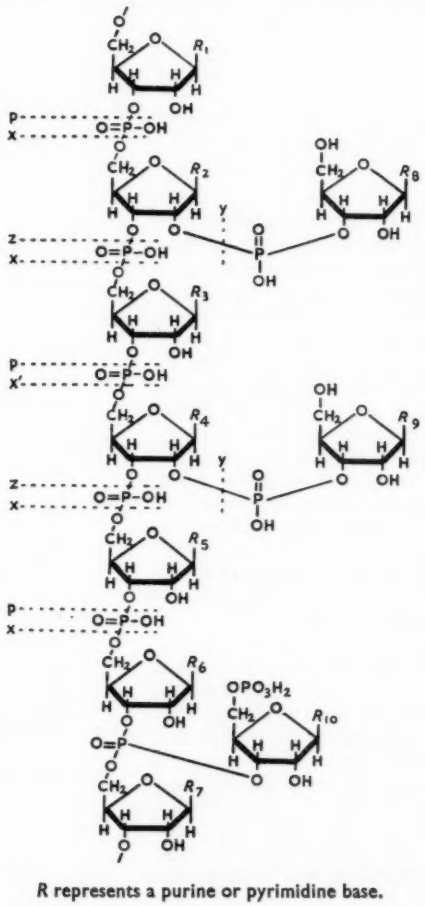
If the trinucleotide illustrated as (I) in fig. 2 is of the type PuPuPy, treatment with the enzyme phosphomonoesterase will remove the terminal phosphate group on $C_{(3')}$ of the lowest nucleotide residue, so that subsequent alkaline hydrolysis will yield two purine nucleotides and a pyrimidine nucleoside (Markham & Smith, 1952). If the same procedure is applied to the polynucleotide shown above containing 16 residues, the terminal residue at the right-hand end of the chain will likewise appear as a pyrimidine nucleoside, which may be separated chromatographically from the mixture of nucleotides also produced. In this way end-groups can be determined qualitatively and quantitatively, and an estimate of the chain length made. If the terminal group consists of a cyclic purine 2′: 3′-phosphate linked at $C_{(5')}$ to a penultimate pyrimidine nucleotide residue, it will be split off by ribonuclease and can be identified chromatographically. Such cyclic purine nucleotides, unlike the corresponding pyrimidine derivatives, are not further acted on by the enzyme.

By methods of this sort Markham & Smith (1952) have identified the cyclic nucleotides adenosine-2′: 3′-phosphate and guanosine-2′: 3′-phosphate, together with adenylic, guanylic, cytidylic and uridylic acids, as end-groups in the polynucleotide chains of the ribonucleic acids of yeast and of turnip yellow mosaic virus.

c. The Occurrence of the $C_{(5')}$-Linkage

The trinucleotide structure illustrated in fig. 2 shows the internucleotide linkages as $C_{(3')}$-O-P-O-$C_{(5')}$. The involvement of $C_{(5')}$ was indicated by the early work of Gulland & Jackson (1938) in Nottingham and has been confirmed by the recent results of Cohn & Volkin (1951), who treated RNA

FIG. 3. A SECTION OF THE POLYNUCLEOTIDE CHAIN IN THE RIBONUCLEIC ACID MOLECULE



R represents a purine or pyrimidine base.

with ribonuclease followed by intestinal phosphatase. The presence of arsenate inhibited any phosphomonoesterase, while not interfering with diesterase activity. The enzymic hydrolysate was found by ion exchange chromatography to contain the 5′-phosphates of all four nucleosides as well as the 3′-phosphates of the pyrimidine nucleosides.

It would appear therefore that $C_{(5')}$ is involved in the main internucleotide linkages which must be either $C_{(2')}$-$C_{(5')}$ or $C_{(3')}$-$C_{(5')}$. Although most authors favour the latter by reason of the analogy with DNA, where $C_{(2')}$ cannot be involved since it carries no hydroxyl group, Markham & Smith (1952) point out that if RNA had such a structure it would tend to be digested by deoxyribonuclease. This does not in fact happen.

A section of a polynucleotide chain in which the main or "backbone" chain consists of $C_{(3')}$-O-P-O-$C_{(5')}$ linkages is shown in fig. 3. Since ribonuclease yields, in addition to larger fragments, only pyrimidine mononucleotides, it has been suggested that these nucleotides are derived from side chains or branches which must occur at frequent intervals in the molecule (Brown & Todd, 1952). Such side chains probably comprise only one nucleotide residue and must occur at $C_{(2')}$, which is linked through the phosphate group to $C_{(3')}$ or $C_{(2')}$ in the first (or only) nucleoside residue in any branch. Only in this way can the alkali lability of the nucleic acid be maintained. Linkage of $C_{(2')}$ in the main chain to $C_{(5')}$ in the side chain would give an alkali-stable structure since cyclization would be impossible.

The structure envisaged by Brown & Todd (1952) is shown in the upper part of fig. 3. This structure has branches consisting of one nucleotide only but longer side chains could be accommodated by extending them by the same $C_{(5')}$-O-P-O-$C_{(3')}$ sequence as in the main chain. Such a structure as this would be converted by dilute alkali into a mixture of mononucleotides by fission at y and then at x, after cyclization as described earlier. If the side chains are pyrimidine nucleoside residues, the first action of ribonuclease would be their removal (probably initially as cyclic phosphates) by fission at the P-O-$C_{(2')}$ linkage at y. Brown and Todd suggest that, while ribonuclease no doubt ruptures some $C_{(3')}$-O-P linkages in the chain, it also brings about rupture of a few of the $C_{(3')}$-O-P linkages as at z. This would produce a fragment (containing R_3 and R_4) bearing at one end a nucleoside residue attached by phosphate through the 5′-position only and therefore oxidizable by periodate by virtue of the free hydroxyl groups at $C_{(2')}$ and $C_{(3')}$. Such fragments are known to occur in the enzymic hydrolysis products. Subsequent treatment of the products of ribonuclease digestion with alkali would yield mononucleotides, but a fragment such as that described above would yield a free nucleoside (containing R_4) and a nucleoside diphosphate (containing R_3) by fission at x′.

Moreover, treatment with ribonuclease followed by phosphodiesterase would yield pyrimidine nucleotides from the side chains and nucleoside 5′-phosphates from the main chain by rupture at p of the $C_{(3')}$-O-P linkages. Such hydrolysis products have in fact been found by Cohn & Volkin (1951, 1952) using intestinal phosphatase or a purified phosphodiesterase from snake venom. When the latter enzyme was employed, the hydrolysis products also included a proportion of purine nucleosides and pyrimidine nucleoside diphosphates. If, as is not unlikely, the top end of the chain shown in fig. 3 carries a purine nucleotide residue and the bottom end a pyrimidine nucleotide residue, hydrolysis by the diesterase at the $C_{(3')}$-O-P linkages as at p would yield a purine nucleoside from the upper terminal group and a pyrimidine nucleoside-3′ (or 2′) : 5′-diphosphate from the lower terminal group.

d. The Occurrence of Branched Chains

Since electrometric titration data (Fletcher, Gulland & Jordan, 1944) indicate the occurrence of secondary phosphoryl dissociations suggesting the presence of adjacent phosphotriester groups, it is probable that side chains appear where phosphotriester groups occur (Jordan, 1952). Such a side chain is depicted at the foot of fig. 3, but the arrangement of the phosphate groups shown in the nucleoside residue carrying R_{10} is only one of several possibilities. The general pattern of structure shown in fig. 3 has been confirmed by Anderson *et al.* (1952), who methylated yeast RNA and subsequently separated by chromatography the methylated riboses liberated by acid hydrolysis.

It should be noted that Markham & Smith (1952) favour a straight, rather than a branched, chain, in which the pyrimidine nucleotides are distributed at random but the purine nucleotides occur with a definite frequency. They have concluded that RNA, as ordinarily prepared, consists of mixtures of several kinds of comparatively short chains, with mean chain lengths of 12 and 53 residues in yeast and turnip yellow mosaic virus ribonucleic acids respectively.

The ultimate action of ribonuclease on RNA is to produce fragments of widely varying molecular weight, together with pyrimidine mononucleotides but no purine mononucleotides. If the enzyme is allowed to act on RNA for a short time, the initial products are the cyclic 2′ : 3′-phosphates of cytidine and uridine together with oligonucleotides or limit polynucleotides terminating in a pyrimidine nucleoside carrying a cyclic phosphoryl group (fig. 4) (Markham & Smith, 1951b, 1952; Brown, Dekker & Todd, 1952). These cyclic phosphates are slowly hydrolysed by ribonuclease to yield cytidylic acid *b* or uridylic acid *b* either as free nucleotides or as the terminal nucleotide residue in the oligonucleotide (fig. 4). (It is of interest to note that the purine-2′ : 3′-phosphates are unaffected by the enzyme.) Subsequent treatment of the oligonucleotides with alkali yields purine *a* and *b* nucleotides by cyclization and fission as described earlier, together with the *b* isomers of the terminal nucleotides. These pyrimidine *b* nucleotides are of course stable in alkaline solution; in acid they quickly yield an equilibrium mixture of the *a* and *b* forms.

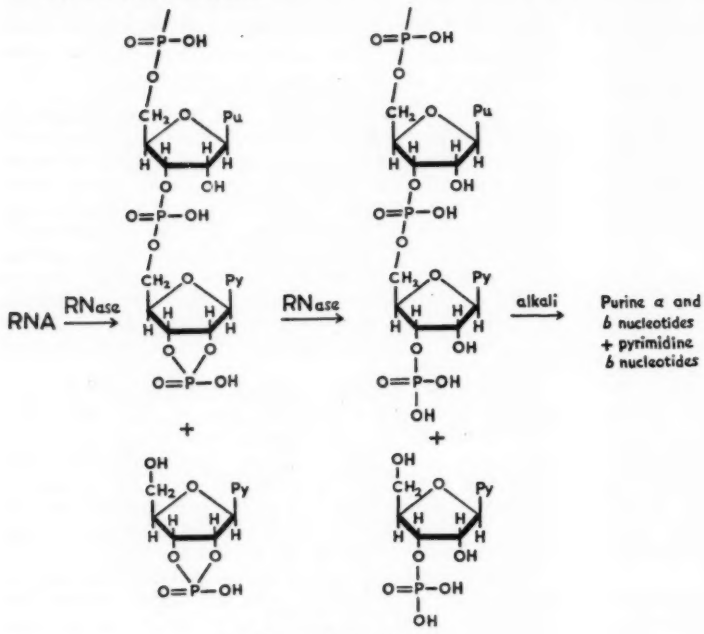
Although the *b* isomers are formulated in fig. 4 as the 3′-phosphates it should be remembered that, as yet, conclusive identification of the *a* and *b* forms as the 2′- and 3′-phosphates respectively is lacking.

4. Deoxyribonucleic Acid (DNA) and Deoxyribonuclease

Deoxyribonucleic acid (DNA) is readily obtained from cell nuclei in the form of highly polymerized fibrous material of which the sodium salt forms a viscous solution showing the phenomenon of streaming birefringence. The molecular weight is of the order of more than a million and the molecules form elongated rods some 8,000 A. long and 10 A. in diameter.

The component parts of DNA include D-2-deoxyribose instead of D-ribose found in RNA, and the same purine bases (Table I). In DNA, however, the pyrimidine base thymine replaces the uracil found in RNA, and deoxyribonucleic acids

FIG. 4. THE ACTION OF RIBONUCLEASE (RNase) ON RIBONUCLEIC ACID (RNA), SHOWING THE INTERMEDIATE FORMATION OF CYCLIC PHOSPHATES



from certain sources have recently been shown by Wyatt (1951) to contain small amounts of 5-methyl cytosine (Table II), while DNA from certain types of bacteriophage contains hydroxy methyl cytosine in place of cytosine (Wyatt & Cohen, 1952). Although the molar proportions of bases vary widely according to the source of the DNA (Table II), the general structural pattern of DNA appears to be that of a polynucleotide formed by the condensation of large numbers of deoxyribonucleotides into a long and probably unbranched chain (fig. 5). It is attacked specifically by the enzyme deoxyribonuclease (DNase).

DNase, like RNase, can be prepared by fractional precipitation with ammonium sulphate of acid extracts of ox pancreas (McCarty, 1946) and has been obtained in crystalline form by Kunitz (1948). It is readily inactivated by heat, is activated by magnesium ions (0.003 M) and inhibited by citrate (0.01 M). Its activity is usually measured by its power to reduce the viscosity of solutions of the sodium salt of DNA.

The mode of action of DNase is less well understood than that of RNase but it consists essentially in the fragmentation of the molecule with the formation of, among other things, at least six dinucleotides and a non-reactive core or "limit polynucleotide." When DNase and intestinal phosphatase act together on DNA, deoxyribonucleotides are obtained which have been identified as the 5′-phosphates. It would appear probable therefore that the internucleotide linkage is $C_{(5')}$-O-P-O-$C_{(3')}$ (fig. 5). As already mentioned, $C_{(2')}$ cannot be involved since it carries no hydroxyl group in the deoxy sugars.

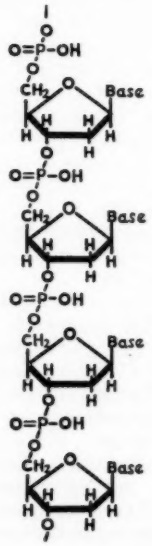
5. Clinical Uses of Deoxyribonuclease

During the last few years deoxyribonuclease has found an interesting application in the liquefaction of viscous purulent material. Tillett and Sherry and their colleagues in America (Sherry, Tillett & Christenson, 1948; Sherry, Johnson & Tillett, 1949; Tillett, Garner & Christenson, 1948; Tillett & Sherry, 1949) have shown that the tough and viscous consistency of many purulent pleural exudates is due largely to deoxyribonucleoprotein derived from the nuclei of pus cells. They have described dramatic clinical improvement following the injection of "streptodornase" (a preparation of DNase from cultures of haemolytic streptococci) into the pleural cavity of patients with purulent and sanguineous pleural exudates.

Armstrong & White (1950) at the Postgraduate Medical School of London have shown that many samples of purulent sputum that are expectorated only with difficulty have a fibrillar structure owing to the presence of deoxyribonucleoprotein. By the administration by spray inhalation of purified DNase prepared from ox pancreas, such sputum may be rendered less viscous and so may be more easily coughed up by patients with chronic bronchial inflammatory diseases. Considerable clinical benefit was obtained with repeated inhalations containing 5 mg. of DNase per day.

In this brief survey it is not possible to discuss the large amount of information which has recently been obtained concerning the biological importance of the

FIG. 5. A SECTION OF THE POLYNUCLEOTIDE CHAIN IN DEOXYRIBONUCLEIC ACID (DNA)



nucleic acids, but it should now be clear that the story of the chemical structure of these complex substances is at last being gradually unfolded. In the attack on this problem the enzymes ribonuclease and deoxyribonuclease have played, and will continue to play, an essential part, but these important enzymes have other applications, not the least of which is the recent use of deoxyribonuclease in the clinical field, and the use of both enzymes as histochemical tools.

References

Anderson, A. S., Barker, G. R., Gulland, J. M. & Lock, M. V. (1952) *J. chem. Soc.*, p. 369

Armstrong, J. B. & White, J. C. (1950) *Lancet*, **2**, 739

Baddiley, J. (1951a) *Rep. Progr. Chem.* **47**, 253

Baddiley, J. (1951b) *Annu. Rev. Biochem.* **20**, 149

Brachet, J. (1940) *C.R. Soc. Biol., Paris*, **133**, 88

Brown, D. M., Dekker, C. A. & Todd, A. R. (1952) *J. chem. Soc.*, p. 2715

Brown, D. M., Magrath, D. I. & Todd, A. R. (1952) *J. chem. Soc.*, p. 2708

Brown, D. M. & Todd, A. R. (1952) *J. chem. Soc.*, p. 52

Chargaff, E. (1950) *Experientia*, **6**, 201

Cohn, W. E. (1951) *J. cell. comp. Physiol.* **38**, Suppl. No. 1, p. 21

Cohn, W. E. & Volkin, E. (1951) *Nature, Lond.* **167**, 483

Cohn, W. E. & Volkin, E. (1952) *Arch. Biochem. Biophys.* **35**, 465

Davidson, J. N. (1950) *The biochemistry of the nucleic acids.* Methuen, London

Davidson, J. N. (1953) *Nucléoprotéines et la croissance des tissus.* In: *II^e Congrès International de Biochimie, Paris, 21-27 Juillet 1952. Compte rendu*, p. 49. Masson, Paris

Davidson, J. N. & Waymouth, C. (1946) *J. Physiol.* **105**, 191

Fletcher, W. E., Gulland, J. M. & Jordan, D. O. (1944) *J. chem. Soc.*, p. 33

Gulland, J. M. & Jackson, E. M. (1938) *J. chem. Soc.*, p. 1492

Jones, W. (1920) *Amer. J. Physiol.* **52**, 203

Jordan, D. O. (1952) *Annu. Rev. Biochem.* **21**, 209

Kunitz, M. (1940) *J. gen. Physiol.* **24**, 15

Kunitz, M. (1948) *Science*, **108**, 19

McCarty, M. (1946) *J. gen. Physiol.* **29**, 123

McDonald, M. R. (1948) *J. gen. Physiol.* **32**, 39

Magasanik, B. & Chargaff, E. (1951) *Biochim. biophys. Acta*, **7**, 396

Markham, R. & Smith, J. D. (1949) *Biochem. J.* **45**, *Proc.* xxxii

Markham, R. & Smith, J. D. (1950) *Biochem. J.* **46**, 513

Markham, R. & Smith, J. D. (1951a) *Biochem. J.* **49**, 401

Markham, R. & Smith, J. D. (1951b) *Nature, Lond.* **168**, 406

Markham, R. & Smith, J. D. (1952) *Biochem. J.* **52**, 552; 558; 565

Sherry, S., Johnson, A. & Tillett, W. S. (1949) *J. clin. Invest.* **28**, 1094

Sherry, S., Tillett, W. S. & Christenson, L. R. (1948) *Proc. Soc. exp. Biol., N.Y.* **68**, 179

Tillett, W. S., Garner, R. L. & Christenson, L. R. (1948) *Proc. Soc. exp. Biol., N.Y.* **68**, 184

Tillett, W. S. & Sherry, S. (1949) *J. clin. Invest.* **28**, 173

Wyatt, G. R. (1951) *Biochem. J.* **48**, 581

Wyatt, G. R. & Cohen, S. S. (1952) *Nature, Lond.* **170**, 1072

* * *

PLASMA ALKALINE PHOSPHATASE IN DISEASE

E. J. KING M.A. Ph.D. D.Sc. F.R.I.C.

Postgraduate Medical School of London

1 Historical
2 Phosphates in fermentation
3 Bone phosphatase
4 Plasma phosphatase (alkaline)
  a Properties
  b Methods of estimation
  c Normal values of alkaline phosphatase
  d Alkaline phosphatase in bone diseases
  e Alkaline phosphatase in jaundice
5 Summary
References

The phosphatases are enzymes which hydrolyse the " organic salts ", or esters, of phosphoric acid, according to the equation: $ROPO_3H_2 + H_2O \rightarrow ROH + H_3PO_4$, where *R*OH is any alcoholic, phenolic, sugar or sugar-like substance, such as ethanol, phenol, glucose or glycerol. There are several phosphatases, in both plant and animal tissues, and they are variously active at acid, neutral and alkaline reactions, and are characteristically activated and inhibited by various substances.

1. Historical

Early in this century the only esterase known was lipase. Fats are esters of glycerol with fatty acids, e.g., glyceryl stearate; and fat-splitting enzymes are esterases in the same sense that a glycerophosphate-splitting enzyme is. Grosser & Husler (1912) described a glycerophosphatase in the kidney; and Plimmer (1913), in studying the occurrence of lipases in animal tissues, was led to the discovery of a phosphate-liberating enzyme in the intestine, which he showed to be active against phosphoric esters of ethanol, hexose and glycerol and against phosphoprotein. Takahishi, however, is usually credited with being the first to describe such an enzyme (Suzuki, Yoshima & Takahishi, 1907). From the preparation known as takadiastase, which is made from a fungus growing on bran, he extracted an enzyme which would actively split simple phosphoric esters into the free alcohol or sugar and inorganic phosphate. Early recognition of similar enzymes, active against several organic phosphates, was made by Harden & Young (1908), McCollum & Hart (1908), Neuberg & Karczaq (1911), Levene & Medigreceanu (1911), Harding (1912) and von Euler & Funke (1912).

2. Phosphates in Fermentation

First recognition of the important role of phosphates in living processes came, however, from the work of Harden & Young (1905) on fermentation. The addition of sodium phosphate accelerated alcohol production, and a fructose-diphosphate, the " Harden ester ", was isolated as a calcium salt from the fermenting solution. The presence in yeast of a sugar-phosphate-synthesizing enzyme was thus demonstrated.

Harden's pupil and co-worker, Robison, found a second phosphoric ester, a sugar monophosphate, whose calcium salt was freely soluble in water, but was insoluble in alcohol (Harden & Robison, 1914; Robison, 1922). This is known as the " Robison ester ".[1]

3. Bone Phosphatase

In studying the action of yeast on his ester, Robison added some of the clear juice expressed from yeast to solutions of his calcium hexosemonophosphate. The mixtures remained entirely clear for some time, but hours later he observed heavy chalk-like precipitates at the bottom of the solutions. These were precipitated calcium phosphate, and Robison was not slow to recognize that a phosphoric ester-splitting enzyme must be present in the yeast juice, which hydrolysed his hexosemonophosphate to free sugar and inorganic phosphate, the latter precipitating as insoluble calcium phosphate. On observing these precipitates and pondering the mechanism of their formation, Robison was struck by the thought that there might be a similar enzyme present in bony tissue; that, indeed, the formation of bone might be brought about by a process which involved a similar mechanism. An extract of young bones gave the same result and obviously contained something capable of hydrolysing hexosephosphate, i.e., a phosphoric ester-splitting enzyme, a phosphatase. This finding led Robison to his well-known theory of bone formation.

Robison (1923) suggested that the bone enzyme in the hypertrophic cells of the tissue where ossification takes place liberates free phosphate from the organic esters of phosphorus contained in the fluids bathing the bone or cartilage, thus giving a local increase in the amount of inorganic phosphate in solution. According to the mass law, any increase in the concentration of phosphate ion would, in the presence of inorganic calcium, lead to a deposition of calcium phosphate.

$$3Ca^{++} + 2PO_4^{---} \rightleftarrows \underset{\text{(in solution)}}{Ca_3(PO_4)_2} \rightleftarrows \underset{\text{(solid)}}{Ca_3(PO_4)_2}$$

When the bones of rachitic rats were immersed in a solution of calcium glycerophosphate a dense deposit of calcium phosphate was formed in the matrix of the hypertrophic cartilage, where ossification normally occurs in the animal receiving a proper diet. Both the seat of action of the enzyme and its ability to bring about calcification were thus shown. That there is a natural substrate for this bone phosphatase in the plasma was shown by Martland & Robison (1924, 1926) and by Kay & Robison (1924). Working with the bones and cartilages of the human foetus, Martland and Robison failed to find any of the enzyme present in cartilage prior to the appearance of ossification centres, but demonstrated its presence wherever ossification was taking place.

In cultures of embryonic bone (rabbit, chick) Fell & Robison (1929, 1930, 1934) demonstrated a parallelism between the amounts of the enzyme and the apparent rates at which calcification was taking place as seen in histological sections. In the dental pulp of embryonic teeth Engel & Furuta (1942) have shown a similar occurrence of the enzyme. On the other hand, the cartilagenous skeletons of the selachian fishes are almost devoid of phosphatase, whereas their teeth are rich

[1] See Dickens, p. 105 of this Bulletin.

in it (Bodansky, Bawkin & Bawkin, 1931; Roche & Bullinger, 1939). Huggins (1931) showed that pieces of bladder wall, transplanted intramuscularly, acquired a high phosphatase activity, if and when heterotopic ossification set in. Botterell & King (1935) found an increase of the enzyme in the sites of fracture repair in the radii of rabbits. In addition to these, and many other, chemical demonstrations of the probably important role of phosphatase, many recent histological studies, with new methods of staining for the enzyme, have confirmed the distribution of phosphatase in the skeleton, and its close association with bone formation and many (though not all) processes of calcification (e.g., Bourne, 1943).

Although Kay & Robison (1924) showed the presence of a natural substrate for the bone phosphatase, the amount of the plasma ester phosphate was very small (Martland & Robison, 1926), and there was some doubt as to whether the quantities of phosphate which its enzymic hydrolysis could liberate would account for all the calcium phosphate deposited in bone. Robison suggested the possibility of a " phosphate cycle " in the bone, whereby phosphoric esters would be synthesized in one stage of a process, and hydrolysed by the phosphatase in the osteoblasts at a later stage. This idea was worked on by Robison & Rosenheim (1934), and has been given detailed expression by Gutman & Gutman (1941) and Roche (1947, 1950), who postulate that a phosphoric ester-synthesizing enzyme (phosphorylase) brings about a phosphorylysis of the glycogen in bone with the inorganic phosphate of the plasma to produce glucose-1-phosphate, which then becomes the principal substrate for the bone phosphatase.

The Robison theory of bone formation is not universally accepted by all workers, and other roles for the bone phosphatase have been suggested (cf. McKelvie, 1948; Siffert, 1951). But it serves as a good working hypothesis which explains most, if not all, of the known facts, and there is no doubt that it has stimulated much useful effort and many fruitful ideas.

4. Plasma Phosphatase (Alkaline)

That the blood plasma also contains a phosphatase, similar in properties to that of bone, was shown by Martland & Robison (1926); and the relative phosphatase activities of the plasma from many clinical conditions were studied by Kay (1929, 1930), who found a marked elevation in generalized bone disease, and normal or nearly normal levels in other diseases.

a. Properties

Effect of pH. The plasma phosphatase, like that of the bone, is markedly dependent on pH. While active at neutrality, and at physiological pH, its activity is much more pronounced at an alkaline reaction. Thus, Robison & Soames (1924) found a very rapid increase in the percentage of ester hydrolysed between pH 7 and 9.4; and at more alkaline reaction the hydrolysis rapidly diminished. The " optimum " pH varies slightly with the phosphate ester used as substrate; ethyl phosphate is optimally hydrolysed at pH 8.1, glucose-phosphate at pH 8.6, glycerophosphate at pH 9.4, phenyl phosphate and nitrophenyl phosphate at pH 9.8 (King & Delory, 1939). Methods of phosphatase determination must take these pH optima into account.

Substrate Specificity. The plasma and bone phosphatases (and those of kidney and intestine) catalyse the hydrolysis of the primary esters only, i.e., the mono-esters of orthophosphoric acid. They do not act on secondary esters (diesters). Monoethyl phosphate is attacked, but diethyl phosphate is not. Nor are they active against the phospholipins and other complex phosphorus-containing organic substances of animal tissues.

Activation and Inhibition. It is characteristic of the alkaline phosphatase—whether it is from plasma, bone, kidney, or intestine—that its activity diminishes as it is dialysed free of magnesium ions; and that the activity is restored on adding a magnesium salt. On the other hand it is made less active by the addition of sulphydryl compounds, such as glutathione, which act as inhibitors.

b. Methods of Estimation

A measure of the phosphatase activity of a tissue or extract could be based on the quantity of calcium phosphate precipitated when the substrate is the calcium salt of a phosphoric ester. In practice, however, something much easier and quicker is required. Robison took advantage of the blue colour of solutions containing reduced phosphomolybdate in order to determine small concentrations of inorganic phosphate and thus developed a quick and convenient way of determining phosphatase activity. A small measured amount of an extract of bone was added to a standard solution of hexosephosphate, whose organically combined phosphate does not react with molybdate; and the amount of inorganic phosphate liberated by enzymic hydrolysis in a certain time was estimated by the intensity of the blue colour produced, which gave the measure of phosphatase activity in the extract.

Glycerophosphate Methods. Kay (1930) used the same reaction for his well-known procedure for plasma phosphatase. Instead of hexosephosphate, he used the more easily obtainable sodium β-glycerophosphate as substrate, and allowed a measured amount of plasma to act on a glycerophosphate solution at 37° C. and at its own pH of about 7.4. This method had the grave disadvantage that the rate of hydrolysis was so slow that only after 48 hours was sufficient inorganic phosphate liberated to give a measurable colour. Although Kay's results on cases of bone disease commanded wide interest, his method of determination took so long that a simpler and quicker procedure was obviously necessary before phosphatase estimations could have any wide clinical application.

Jenner & Kay (1932) modified Kay's method by conducting the hydrolysis at a pH nearer the optimum (pH 8.8, glycine buffer), thus gaining much more analysable inorganic phosphate in a much shorter time. By the use of a very sensitive method for phosphate (molybdate-stannous chloride) they were able to cut the hydrolysis time to 3 hours. Plasma (0.25 ml.) is allowed to act on the buffered glycerophosphate at 37° C., the hydrolysis is stopped with trichloroacetic acid, and the increase of inorganic phosphate (there is a small amount initially present in the plasma) is measured. Their unit of phosphatase activity is defined as that amount of enzyme which will liberate inorganic phosphate corresponding to 1 mg. of P from excess glycerophosphate at 37° C. and pH 8.8 in the 3 hours; and the units per 100 ml. of plasma are equal to the number of milligrams of inorganic P which are hydrolysed under these conditions. By this method average normal plasmas give about 6 units, with a range of

3 to 13 units. The King and Armstrong method (see below) yields similar values (King & Armstrong, 1934).

Bodansky (1933) has described a rather similar method for phosphatase, which is much used, particularly in America. The hydrolysis, of 0.5% glycerophosphate, is carried out in a barbitone buffer (pH 8.6), and the units are equal to the number of milligrams of P liberated in 1 hour. Although it is not possible to equate exactly one unit to another, that of Jenner and Kay is on the average about 2½ times the Bodansky unit, e.g., a plasma having 10 Jenner and Kay units would show about 4 units by the Bodansky method.

Shinowara, Jones & Reinhart (1942), and others, have described similar methods, based on glycerophosphate.

Phenyl-Phosphate Methods. King & Armstrong (1934), in seeking a quicker and more convenient procedure for plasma phosphatase determination, took advantage of the fact that the phosphoric ester of phenol is hydrolysed twice as quickly as glycerophosphate. Moreover, there was a colorimetric method for phenol (Folin & Ciocalteu, 1927) which was quicker and more sensitive than existing methods for phosphate. Since, in the hydrolysis of phenyl phosphate, a molecule of phenol is liberated for each molecule of inorganic phosphate, the rate of hydrolysis can be determined as satisfactorily by estimating one as the other. Because of the faster hydrolysis and the sensitive phenol method, King and Armstrong were able to reduce both the amount of plasma used and the time of hydrolysis. In their original method, in 1934, the incubation time was 30 minutes and a veronal buffer (pH 9.4) was used. By using the exact pH optimum (pH 9.8 at 37° C.), a more recent modification has reduced the time of hydrolysis to 15 minutes (King, 1946, 1951), thus increasing still further the speed and ease of the determination, as well as its accuracy, since the greatly decreased time of hydrolysis yields results approximating to initial velocities. Plasma (0.2 ml.) is mixed with 4 ml. of a solution of phenyl phosphate in sodium carbonate-bicarbonate buffer, kept at 37° C. for 15 minutes, the enzyme action stopped and the proteins precipitated by the addition of Folin and Ciocalteu's phenol reagent, and the blue colour proportional to free phenol is developed in the filtrate by the addition of sodium carbonate. The units are equal to the number of milligrams of phenol ydrolysed per 100 ml. of plasma. They are equal to the Jenner and Kay units, and are on the average 2½ times the Bodansky units. Several modifications of this method have been described, e.g., by Buch & Buch (1939), Greenburg (1940) and Aebi (1952).

p-*Nitrophenyl-Phosphate Methods.* Ohmori (1937) was the first to use this substrate for phosphatase estimation; and King and Delory again used it in 1939. It is attacked even faster by the enzyme than is phenyl phosphate (King & Delory, 1939, 1943). The *p*-nitrophenol which is liberated is of an intense yellow colour at an alkaline reaction and can be determined directly without adding any reagent other than strong alkali, which serves both to stop the enzyme action and to develop the yellow colour of the nitrophenol. Bessey, Lowry & Brock (1946) have based a plasma phosphatase method on this reaction.

Phthalein-Phosphate Methods. Lawford and King prepared the phosphoric ester of phenolphthalein, which is colourless in alkaline solution because its potential quinonoid structure is blocked by the phosphate group (cf. Lawford, 1937; King, 1938). On hydrolysis by phosphatase, free phenolphthalein is liberated and the solution slowly turns pink. King (1943) described several phthalein and sulphonephthalein phosphates (phenol- and thymol-phthalein, phenol red, thymol blue, bromophenol blue), but thought that none of them made as useful a substrate for phosphatase determination as the phenyl phosphate of King and Armstrong, although Bray & King (1943) found phenolphthalein diphosphate to be a useful reagent for differentiating bacteria. Huggins & Talalay (1945), however, have considered that it furnishes a suitable and convenient method for plasma phosphatase estimation, and have based an assay procedure on the amount of phenolphthalein liberated under defined conditions of pH, temperature, time, etc.

Besides the phenolphthalein diphosphate, a phenolphthalein monophosphate has been made by Gutman (1950) and by King in 1948 (unpublished work). It is likewise colourless, and its solution turns pink when hydrolysed by phosphatase in an alkaline buffer. We have used this substance for plasma phosphatase estimations, but regard it also as inferior for the purpose to phenyl phosphate. Moreover, the several preparations of phenolphthalein monophosphate which have been made have differed in their rates of phosphatase hydrolysis, and it does not appear that the method of synthesis used yields a single chemical substance.

c. Normal Values of Alkaline Phosphatase

The two most commonly used methods for plasma phosphatase are those of Bodansky (1933) and of King & Armstrong (1934) described above. The figures given for plasma phosphatase levels in health and disease, in the following sections of this paper, are in the units of King and Armstrong, who gave the normal range as 3–13 units for adults, and up to 20 units for children (cf. Buch & Buch, 1939). Young *et al.* (1946) studied plasma phosphatase levels in healthy women in pregnancy in the United Kingdom, and Stern (1948) in children and adults in Germany. These representative findings are summarized in Table I.

d. Alkaline Phosphatase in Bone Diseases

Rickets. Many authors have noted an elevation, often marked, of the plasma phosphatase in clinical rickets (Morris & Peden, 1937; Patwardhan, Chitre & Sukhatankar, 1944; Cantarow & Trumper, 1945). A good summary is given by Gray & Carter (1949), and their own findings (cf. Table II)

TABLE I. NORMAL PLASMA PHOSPHATASE

Subjects	King-Armstrong phosphatase units
Healthy laboratory workers and others (British, Canadian, Danish): 200 cases	7.8 (3–13)
Pregnant women (British)	
Stage of pregnancy: first 24 weeks: 261 cases	4.3 ± 2.0
32 weeks to term: 220 cases	10.5 ± 3.0
post partum (1 week)	10.4 ± 3.5
post partum (6 months)	7.6 ± 2.5
German civilians	
35 men	4.0 ± 2.0
33 women	3.0 ± 2.1
71 children	12.5 (5–28)

TABLE II. PLASMA PHOSPHATASE IN NORMAL AND RACHITIC INFANTS FROM ONE MONTH TO THREE YEARS OLD (Gray & Carter, 1949)

Number of infants examined	Diagnosis	King-Armstrong phosphatase units	
		Average	Range
56	Normal controls	17	11–20
64	Clinical rickets	25–40*	25–76*
9	Early rickets	35*	—

* The values fell to normal on treatment, or 3 weeks after administration of 200,000 units of calciferol.

are a very helpful guide to the use of these estimations in the diagnosis and treatment of this disease.

Osteomalacia. This adult form of rickets also yields enhanced values, usually 20–30 King-Armstrong units, but often higher.

Osteitis Fibrosa Cystica. A high plasma phosphatase is almost invariable in von Recklinghausen's disease. This, with the high serum calcium, is a reliable diagnostic finding. Depending on the state of development of the disease, values up to 200 units (occasionally over) may be found, but the usual range is 40–100 units.

Osteitis Deformans. Paget's disease likewise shows a similarly elevated plasma phosphatase, but the nearly normal serum calcium serves to give a useful contrast to von Recklinghausen's disease, in which phosphatase and calcium are both high. Paget's disease shows considerable variability in the phosphatase values during the course of the disease, as is shown in Table III.

Fractures. Although there is a big local increase in the bone phosphatase at the site of a healing fracture (Botterell & King, 1935), this is only mildly reflected in the plasma, which usually has 10–20 units, or a little more occasionally (cf. Peden, 1937).

Bone Tumours. In general it may be said that any case of bone cancer, in which active replacement is going on, accompanied by osteoblastic activity, will show a raised plasma phosphatase, e.g., in osteogenic sarcoma, and metastatic carcinoma; but those malignant conditions of bone (multiple myeloma and Ewing's tumour) which are destructive without replacement, and which show high osteoclastic activity, do not cause excess phosphatase to be made, and no elevation of the enzyme in the plasma occurs; see, for instance, Franseen, Simmons & McLean (1939), Woodward (1942), Sullivan, Gutman & Gutman (1942), Sunderman (1942) and King & Delory (1948). A summary of the biochemical findings in bone diseases is given in Table IV.

TABLE III. VARIATION OF PHOSPHATASE IN PAGET'S DISEASE

Case number	Periodic determinations (King-Armstrong phosphatase units)						Period of observation
1	41	52	23	65	—	—	3 months
2	52	62	36	75	67	70	4 years
3	39	54	45	45	66	—	3 months
4	81	—	79	—	—	—	4 months
5	33	29	37	32	34	30	1 year

e. Alkaline Phosphatase in Jaundice

Roberts (1933) was the first to show that bone diseases were not unique in causing a high plasma phosphatase. He found gross elevation in obstructive jaundice, and nearly normal levels in non-obstructive cases. Armstrong, King & Harris (1934) brought experimental proof of this finding by ligating, or otherwise obstructing, the common bile ducts of dogs, thereby causing enormous increases (up to 500 units) in the plasma phosphatase, as well as in the bile pigments. On relief of the obstruction, the jaundice subsided and the phosphatase slowly returned to normal levels. Similar results were obtained by Bodansky & Jaffe (1934), and have since been confirmed by numerous workers, e.g., Freeman, Chen & Ivy (1938), Gutman, Hogg & Olson (1940), Dalgaard (1947, 1949) and Jackson (1952). In contrast to the very high levels encountered in experimental obstructive jaundice, only moderately elevated values were found in toxic jaundice produced by chloroform and by phosphorus poisoning of the liver; and experimental haemolytic jaundice (hydrazine, toluidine) caused no rise of the plasma phosphatase (Armstrong & King, 1935).

Armstrong *et al.* (1934) described several cases of clinical obstructive jaundice, due, for example, to carcinoma of the head of the pancreas, in which markedly raised plasma phosphatase values were found (> 100 units); after operation there was a gradual decline until normal levels were reached. Herbert (1935) made an extensive study of the enzyme in the different types of jaundice, as did Maclagan (1944) and Sherlock (1946). Their findings, and our own numerous results, lead to the following summarized conclusions: in obstructive jaundice in which the obstruction is nearly complete or of long standing, values of over 30 and up to 200 units commonly occur, while less than 30 units may be found if the obstruction is partial or intermittent; in infective hepatitis, cirrhosis and toxic jaundice about 15–30 units are probable; while the plasma of haemolytic jaundice cases gives phosphatase values within the normal range.

Similar findings, though expressed in different units, have been reported by Greene (1934), Meranze, Meranze & Rathman (1939), Gutman, Olson, Gutman & Flood (1940), Bodansky & Jaffe (1934) and Jaffe & Bodansky (1943) (see Cantarow & Trumper, 1949). Sherlock (1946) gives valuable tables of the findings for plasma phosphatase and other biochemical estimations, and describes the use that can be made of them in the diagnosis and assessment of liver disease.

TABLE IV. BIOCHEMICAL FINDINGS IN BONE DISEASES

Condition	Calcium	Phosphorus	Phosphatase
Rickets	Normal	Low	High
Osteomalacia	Normal	Low	High
Tetany	Low	Normal	Normal
Osteitis fibrosa cystica	High	Normal	High
Osteitis deformans	Normal	Normal	High
Metastatic carcinoma	Normal	Normal	Normal (to high)
Osteogenic sarcoma	Normal	Normal	High
Multiple myeloma	High (often)	Normal	Normal
Fractures	Normal	Normal	High (slightly)

5. Summary

The presence of alkaline phosphatase in blood plasma and the abnormalities of concentration of this enzyme encountered in disease, together with methods for its determination, have been described.

Plasma phosphatase is higher in children than in adults. Elevated values occur in rickets, osteomalacia, osteitis fibrosa cystica, osteitis deformans, certain bone tumours and in healing fractures.

In jaundice very high plasma phosphatase values may be found in cases of nearly complete or prolonged obstruction, less high when the obstruction is incomplete or intermittent, moderately raised in hepatitis and cirrhosis, and normal in haemolytic jaundice.

The determination of plasma phosphatase is a useful aid in the diagnosis of generalized bone disease and the different kinds of jaundice.

References

Aebi, H. (1952) *Schweiz. med. Wschr.* **82,** 135
Armstrong, A. R. & King, E. J. (1935) *Canad. med. Ass. J.* **32,** 379
Armstrong, A. R., King, E. J. & Harris, R. I. (1934) *Canad. med. Ass. J.* **31,** 14
Bessey, O. A., Lowry, O. H. & Brock, M. J. (1946) *J. biol. Chem.* **164,** 321
Bodansky, A. (1933) *J. biol. Chem.* **101,** 93
Bodansky, A., Bawkin, R. M. & Bawkin, H. (1931) *J. biol. Chem.* **94,** 551
Bodansky, A. & Jaffe, H. L. (1934) *Proc. Soc. exp. Biol. N.Y.* **31,** 107
Botterell, E. H. & King, E. J. (1935) *Lancet,* **1,** 1267
Bourne, G. (1943) *Quart. J. exp. Physiol.* **32,** 1
Bray, J. & King, E. J. (1943) *J. Path. Bact.* **55,** 315
Buch, I. & Buch, H. (1939) *Acta med. scand.* **101,** 211
Cantarow, A. & Trumper, M. (1945) *Clinical biochemistry,* 3rd ed. Saunders, London
Cantarow, A. & Trumper, M. (1949) *Clinical biochemistry,* 4th ed. Saunders, London
Dalgaard, J. B. (1947) *Acta physiol. scand.* **13,** 310
Dalgaard, J. B. (1949) *Acta physiol. scand.* **16,** 287
Engel, M. B. & Furuta, W. (1942) *Proc. Soc. exp. Biol., N.Y.* **50,** 5
Euler, H. von & Funke, Y. (1912) *Z. phys. Chem.* **77,** 488
Fell, H. B. & Robison, R. (1929) *Biochem. J.* **23,** 767
Fell, H. B. & Robison, R. (1930) *Biochem. J.* **24,** 1905
Fell, H. B. & Robison, R. (1934) *Biochem. J.* **28,** 2243
Folin, O. & Ciocalteu, V. (1927) *J. biol. Chem.* **73,** 627
Franseen, C. C., Simmons, C. C. & McLean, R. (1939) *Surg. Gynec. Obstet.* **68,** 1038
Freeman, S., Chen, Y. P. & Ivy, A. C. (1938) *J. biol. Chem.* **124,** 79
Gray, J. D. & Carter, F. S. (1949) *Arch. Dis. Childh.* **24,** 189
Greenburg, D. (1940) *J. Lab. clin. Med.* **25,** 634
Greene, C. H. (1934) *J. clin. Invest.* **13,** 1079
Grosser, P. & Husler, J. (1912) *Biochem. Z.* **39,** 1
Gutman, A. B. (1950) *Fed. Proc.* **9,** 180
Gutman, A. B. & Gutman, E. B. (1941) *Proc. Soc. exp. Biol., N.Y.* **48,** 687
Gutman, A. B., Hogg, B. M. & Olson, K. B. (1940) *Proc. Soc. exp. Biol., N.Y.* **44,** 613
Gutman, A. B., Olson, K. B., Gutman, E. B. & Flood, C. A. (1940) *J. clin. Invest.* **19,** 129
Harden, A. & Robison, R. (1914) *Proc. chem. Soc., Lond.* **30,** 16
Harden, A. & Young, W. J. (1905) *Proc. chem. Soc., Lond.* **21,** 189
Harden, A. & Young, W. J. (1908) *Proc. roy. Soc.* B, **80,** 299
Harding, V. J. (1912) *Proc. roy. Soc.* B, **85,** 418
Herbert, F. K. (1935) *Brit. J. exp. Path.* **16,** 365
Huggins, C. B. (1931) *Biochem. J.* **25,** 728
Huggins, C. B. & Talalay, P. (1945) *J. biol. Chem.* **159,** 399
Jackson, S. H. (1952) *J. biol. Chem.* **198,** 553
Jaffe, H. L. & Bodansky, A. (1943) *Bull. N.Y. Acad. Med.* **19,** 831
Jenner, H. D. & Kay, H. D. (1932) *Brit. J. exp. Path.* **13,** 22
Kay, H. D. (1929) *Brit. J. exp. Path.* **10,** 253
Kay, H. D. (1930) *J. biol. Chem.* **89,** 235; 249
Kay, H. D. & Robison, R. (1924) *Biochem. J.* **18,** 755
King, E. J. (1938) *J. Soc. chem. Ind., Lond.* **57,** 85
King, E. J. (1943) *J. Path. Bact.* **55,** 311
King, E. J. (1946) *Micro-analysis in medical biochemistry.* Churchill, London
King, E. J. (1951) *Micro-analysis in medical biochemistry,* 2nd ed. Churchill, London
King, E. J. & Armstrong, A. R. (1934) *Canad. med. Ass. J.* **31,** 376
King, E. J. & Delory, G. E. (1939) *Biochem. J.* **33,** 1157; 1185
King, E. J. & Delory, G. E. (1943) *Biochem. J.* **37,** 547
King, E. J. & Delory, G. E. (1948) *Postgrad. med. J.* **24,** 299
Lawford, F. H. (1937) *The enzymatic hydrolysis of phosphoric esters.* (Thesis for Ph.D. degree.) University of Toronto, Toronto
Levene, P. A. & Medigreceanu, F. (1911) *J. biol. Chem.* **9,** 65; 375; 389
McCollum, E. V. & Hart, E. B. (1908) *J. biol. Chem.* **4,** 497
McKelvie, A. M. (1948) *Proc. Mayo Clin.* **23,** 449
Maclagan, N. F. (1944) *Brit. med. J.* **2,** 363
Martland, M. & Robison, R. (1924) *Biochem. J.* **18,** 1354
Martland, M. & Robison, R. (1926) *Biochem. J.* **20,** 847
Meranze, T., Meranze, D. R. & Rathman, M. M. (1939) *Rev. Gastroenterol. Méx.* **6,** 254
Morris, N. & Peden, O. E. (1937) *Quart. J. Med.* **6,** 211
Neuberg, C. & Karczaq, L. (1911) *Biochem. Z.* **36,** 60
Ohmori, Y. (1937) *Enzymologia,* **4,** 217
Patwardhan, V. N., Chitre, R. G. & Sukhatankar, D. R. (1944) *Indian J. med. Res.* **32,** 31
Peden, O. D. (1937) *Arch. Dis. Childh.* **12,** 87
Plimmer, R. H. A. (1913) *Biochem. J.* **7,** 43
Roberts, W. M. (1933) *Brit. med. J.* **1,** 734
Robison, R. (1922) *Biochem. J.* **16,** 809
Robison, R. (1923) *Biochem. J.* **17,** 286
Robison, R. & Rosenheim, A. H. (1934) *Biochem. J.* **28,** 684
Robison, R. & Soames, K. M. (1924) *Biochem. J.* **18,** 740
Roche, J. (1947) *Ann. Nutr., Paris,* **1,** 3
Roche, J. (1950) In: Sumner, J. B. & Myrbäck, K. D. R., ed. *The enzymes: chemistry and mechanism of action,* vol. 1, pt 1, p. 503. Academic Press, New York
Roche, J. & Bullinger, E. (1939) *Bull. Soc. Chim. biol., Paris,* **21,** 166
Sherlock, S. P. V. (1946) *J. Path. Bact.* **58,** 523
Shinowara, G. Y., Jones, E. M. & Reinhart, H. L. (1942) *J. biol. Chem.* **142,** 921
Siffert, R. S. (1951) *J. exp. Med.* **93,** 415
Stern, M. I. (1948) *Brit. J. Nutr.* **1,** 182
Sullivan, T. J., Gutman, E. B. & Gutman, A. B. (1942) *J. Urol.* **48,** 426
Sunderman, F. W. (1942) *Amer. J. clin. Path.* **12,** 404
Suzuki, U., Yoshima, K. & Takahishi, M. (1907) *Bull. Coll. Agric. Tokyo,* **7,** 503
Woodward, H. Q. (1942) *Cancer Res.* **20,** 497
Young, J., King, E. J., Wood, E. & Wootton, I. D. P. (1946) *J. Obstet. Gynaec., Brit. Emp.* **53,** 251

Book Reviews

The Metabolism of Protein Constituents in the Mammalian Body

S. J. Bach. London: Geoffrey Cumberlege, 1952. viii + 272 pages. 24 × 16 cm. £2

Our knowledge of the metabolism of nitrogenous compounds, particularly of amino-acids, has been greatly expanded during the last fifteen or twenty years. This has been due partly to the efforts of many workers investigating pure enzyme or tissue systems prepared in vitro, but perhaps even more important has been the introduction of isotopic tracers. In addition, the use of mixtures of pure amino-acids in feeding experiments, which was made possible by the work of W. C. Rose, has been of great assistance in the field of intermediary metabolism of amino-acids and related compounds. The use of these varied techniques has produced a large volume of work, and has completely changed most of our traditional concepts. This large amount of information is contained in innumerable original articles in various biochemical journals and an urgent need has been felt by all biochemists for a comprehensive and critical review of this rather large subject. It is therefore to be welcomed that this difficult task has been undertaken by Dr S. J. Bach.

The book under review deals with the metabolism of glycine, serine, threonine, valine, leucine, isoleucine, cystine and methionine. The aromatic, acidic and basic amino-acids are not treated in the present volume but it is to be hoped that the author will complete his task and give us a review, equally competent and complete, of the remaining protein constituents. It is particularly useful that the author has included a detailed and up-to-date discussion of related metabolic fields such as the wider aspects of transmethylation, biosynthesis of porphyrins and of uric acid. A particularly valuable feature is that each section is prefaced by a short historical account, and the older literature is fully discussed. As Dr Bach himself says, he has concentrated on conveying to the reader his own interpretation and this certainly lends additional interest to the book. Workers in the field of metabolism will not always fully agree with Dr Bach's conclusions but the evidence is always fairly and fully presented and, so far as the reviewer is aware, no important facts have been withheld. The general production and the printing are up to the high standard expected from the publishers, and the proofreading appears to have been equally good.

All biochemists in the field of protein metabolism are greatly indebted to Dr Bach for this useful and critical review and it might therefore seem somewhat ungracious to point out mistakes or points on which the reviewer disagrees with the author. But it would seem that the statement that norvaline is a protein constituent (p. 89) is not really justified. This amino-acid was claimed to have been isolated by Abderhalden but, so far as the reviewer is aware, this observation has not been confirmed by modern methods. Equally, no evidence or reference is given for the statement that isocysteine or isocystine "has only recently been discovered as a component of the protein molecule..." (p. 148). We should also question whether it is correct to state (p. 60) that "In the chick the mechanism responsible for the synthesis of glycine is absent." What nutritional experiments have shown is that the rate of synthesis of glycine by the animal is not sufficient to cover the requirements of optimum growth. This is understandable if it is remembered that birds require much larger amounts of glycine than mammals since they excrete nitrogen in the form of uric acid, a compound for which glycine is a precursor.

These relatively minor criticisms do not detract from the great value of this book, which is a complete and up-to-date account of an important field of biochemistry. It can be warmly recommended to the specialist in the field and to the advanced student of biochemistry.

A. Neuberger

The Microbiological Assay of the Vitamin B-Complex and Amino Acids

E. C. Barton-Wright. London: Sir Isaac Pitman & Sons, Ltd., 1952. x + 179 pages; 25 figures. 22 × 14 cm. 18s.

There is no need to stress the importance of microbiological methods for the quantitative determination of the B-group vitamins and amino-acids. In most cases such techniques enable estimation of these substances when these are present in concentrations far below the limit of the most sensitive chemical methods, and they are in general fairly specific. Dr Barton-Wright has had wide experience in this field and has provided in this book full details of techniques and methods of computation that he has found satisfactory; many references to other methods are also included. Detailed descriptions are given for the maintenance of stock cultures of the assay organisms and for the preparation of the various types of media and solutions.

The author has dealt with the subject from a strictly practical point of view and has emphasized the necessity for keeping rigidly to the described procedures. In a book of this size it is impossible to deal in any detail with the theoretical background and with the many difficulties that may arise in the assay of substances in natural materials. Attention has to be given to the possible presence of compounds other than that under assay, which may interfere specifically with the determination. The author points out the inhibitory interrelationships which have been observed with amino-acids (and other compounds also) where an excess of one substance may interfere with the utilization of another. Many instances are also known of the elimination or modification of growth-factor requirements by other substances. The assay organism, for instance, may respond equally well to precursors of the factor under examination. Thus, anthranilic acid or indole can interfere with the assay for tryptophan using *Lactobacillus arabinosus*, since this organism can convert these compounds to the amino-acid; the author recommends that samples should be shaken with ether to remove these precursors. The requirement for a B-group vitamin may also be abolished by the presence of a compound of unrelated chemical structure, which may be a product of an enzymic reaction requiring the vitamin (or a higher form) as co-factor. Purines and thymine, for example, promote growth of *Streptococcus faecalis* (recommended for the assay of folic acid) in a medium devoid of folic acid; they can similarly interfere with the assay for *p*-aminobenzoic acid by mutant strains of *Neurospora*, since they are formed by mechanisms in which folic acid (derived from *p*-aminobenzoic acid) participates as co-factor. The requirement of micro-organisms for amino-acids may be influenced considerably by the composition of the test medium, as well as by substances which may be added in the test sample. For instance, Dr Barton-Wright has recommended *Leuconostoc mesenteroides* for the assay of serine; this organism is, however, unsuitable for the assay of serine in samples containing *Leuconostoc citrovorum* factor, since it can convert glycine (present in the basal medium) to serine when the factor is present. Since the substance is acid-labile, it does not interfere with the assay of serine in acid hydrolysates. It is clear that detailed knowledge of the nutrition of an organism must be obtained before it can be used successfully for assay purposes.

Dr Barton-Wright does not mention *Leuconostoc citrovorum* factor (5-formyl-5 : 6 : 7 : 8-tetrahydropteroylglutamic acid) in his discussion of the folic acid group; this compound appears to have greater biological significance than pteroylglutamic acid and it is also active for *Streptococcus faecalis*. A method of assay for vitamin B_{12} has not been included in the book; methods are, however, given for the assay of the other well-known vitamins of the B group, together with inositol and choline, and for 18 amino-acids.

This book will undoubtedly be very valuable to those engaged upon the assay of these substances in food products and related materials.

June Lascelles

Ciba Foundation Colloquia on Endocrinology. Volume I. Steroid Hormones and Tumour Growth, and Steroid Hormones and Enzymes

Ciba Foundation. Edited by *G. E. W. Wolstenholme*, assisted by *Margaret P. Cameron*. London: J. & A. Churchill Ltd., 1952. xx + 315 pages; 48 illustrations. 21 × 14 cm. £1 10s.

This carefully produced volume is the first in a series devoted to the proceedings of colloquia on selected topics in endocrinology, held in London by the Ciba Foundation and attended by invitation by research workers from several countries. Two colloquia are dealt with in each volume; the titles in the present one are *Steroid hormones and tumour growth* and *Steroid hormones and enzymes*.

The papers read at the colloquium on tumour growth are given at varying length and each is followed by a verbatim account of the ensuing discussion. This was usually lively and one receives remarkably vivid impressions (possibly mistaken) of the personalities of some of the participants. As the chairman (Professor A. Haddow) points out, it is difficult to summarize the proceedings of this meeting. In any case, the present reviewer is not qualified to do so. Nobody, however, can fail to be impressed by the promise held out by the work described, of leading to some understanding of the causes of cancer.

Papers presented at the colloquium on steroid hormones and enzymes are reproduced only in summary, but the discussions are again given in full. As noted by the chairman (Dr S. J. Folley), there was a glaring omission from the subjects discussed, viz., the action of steroid hormones on isolated enzymes, for the simple reason that no information is yet available on this subject. Several papers were concerned with the effects in vivo of steroid hormones on the activities of various enzyme systems in various tissues. Most of the remaining papers dealt with β-glucuronidase, an enzyme which there is reason to believe may play some part in the metabolism and action of steroid hormones. Various controversial points on β-glucuronidase were discussed. It was generally agreed that it is unlikely that this enzyme is responsible for the synthesis of conjugated glucuronides, but that it could conceivably transfer glucuronic acid from one compound to another. Discrepancies between results obtained by different workers in attempting to establish a relationship between the β-glucuronidase activity of a tissue and its state of proliferation were, it was suggested, due to strain and species differences in the experimental animals.

In a paper by L. A. Elson, the following theory was put forward to explain the connexion between β-glucuronidase and tissue growth. It is assumed that a hypothetical steroid growth hormone circulates in the body as an inactive glucuronide and is liberated locally in the active form as required, this process being controlled by variations in tissue β-glucuronidase activity. This is the first appearance in print of a theory to which we were attracted when, in Edinburgh, we were first able to suggest that there is a relationship of some kind between β-glucuronidase activity and growth. Among ourselves, it was known as the " proliferogen " theory. Unfortunately, there is no doubt that, at least sometimes, the enzyme is under the control of steroid hormones, rather than the other way round. The theory would require the further postulate that in the uterus, for example, oestrogens do not themselves stimulate growth, but, as a result of their action on the β-glucuronidase activity of the tissue, cause the liberation of the less specific " proliferogen " in relatively high concentration. At this rate the theory would far outstrip the available evidence.

The organizers of these colloquia are to be congratulated on creating an atmosphere of informal speculation. One wonders, however, whether the publication of the verbatim record may not have a restraining effect on future meetings. In any case, the discussions as they stand do not merit publication in such a permanent and expensive form. They do not compete with the usual symposium or review since they cannot be relied upon for a balanced or comprehensive account of their subject. It is unlikely that much of the work of any importance that was new at the date of the colloquia (1950) has not been published since in greater detail through the ordinary channels. These remarks are certainly true of the material on β-glucuronidase.

G. A. Levvy

Symposia of the Society for Experimental Biology. Number VI. Structural Aspects of Cell Physiology

Society for Experimental Biology. Cambridge: University Press, 1952. vii + 357 pages; plates; figures. 25 × 16 cm. £2 5s.

It is the custom of the Society for Experimental Biology from time to time to hold a symposium lasting several days for the discussion of some topic of current biological interest and to invite a group of speakers actively engaged in research in the field under discussion to contribute papers on their recent work. Each paper is presented and discussed in an atmosphere reasonably free from the sense of hurry and strain which so often now attaches to scientific meetings.

The present volume is the sixth of the series of such symposia to be published and contains the papers presented at Bristol in July 1951, on the general subject of *Structural aspects of cell physiology*. As might be expected, the contents cover a wide range of topics, from the physical chemistry of permeability and secretion (Danielli) and nerve conduction (Katz) to virus reproduction (Sanders), genetics (Pontecorvo) and problems of mitosis (Swann, Mitchison, Hughes).

While most of the 21 papers are by British authors, the list includes such recognized authorities from the Continent and the USA as Runnström (the cell surface), Frey-Wyssling (plant cell walls), Lindgren (yeast), Coons (antigens) and Brachet, who contributes an important survey of his recent views on the role of the nucleus in synthesis and morphogenesis.

Some of the papers deal with experimental work of the highest quality which has not hitherto been published in detail, such as Callan's account of the separation of the components of amphibian oocyte nuclei. Others deal with pieces of work which, although manifestly incomplete at the time of the symposium, were nevertheless ripe for discussion; while some, such as the contribution by Walker & Yates, have been published in a more complete form in the rather long interval between the holding of the symposium and the appearance of the published account.

The book is excellently produced and well illustrated, the beautiful electron micrographs in the papers of Manton, Frey-Wyssling and Preston being especially noteworthy. There is an abundance of stimulating material in it to hold the interest of all the various types of biologist—cytologists, geneticists, biophysicists, biochemists and others—who follow modern developments in cell physiology.

J. N. Davidson

The Bile Pigments

C. H. Gray. London: Methuen & Co. Ltd.; New York: John Wiley & Sons, Inc., 1953. (Methuen's Monographs on Biochemical Subjects.) xv + 142 pages; figures. 17 × 11 cm. 9s. 6d.

This latest addition to the series of Monographs on Biochemical Subjects, edited for Methuen by Sir Rudolph Peters and Professor F. G. Young, will appeal to a wide range of readers. Previously only highly specialized, and correspondingly expensive, works on the bile pigments have been available; but the present small volume provides a brief, clear and accurate guide to the subject, and should prove of value to students of biochemistry, physiology and medicine, as well as to medical men, pathologists and chemists who require a concise general account. Medical readers will find excellent chapters on jaundice, on the transport and excretion of bile pigments, and on their qualitative and quantitative detection and estimation. These analytical methods are given in sufficient detail to enable them to be followed in the laboratory from the descriptions.

The theoretical sections of the book, which deal with chemical structure, formation by catabolism of haemoglobin, transport in the blood, and biosynthesis of the bile pigments, are surprisingly comprehensive, considering the restricted space at the author's disposal, and are also fully up to date. The book as a whole is clearly written and forms a very readable account of its subject. The numerous references to original sources at the end of each chapter are well chosen and very usefully supplement the text.

F. Dickens

Dynamic Aspects of Biochemistry

Ernest Baldwin. 2nd ed. Cambridge: University Press, 1952. xx + 544 pages; 38 figures. 22 × 14 cm. £1 10s.

The first duty of a reviewer dealing with a new textbook is to inform the reader about its scope and contents. A new edition of a book with an established place in the pedagogic landscape confronts him with a sterner task.

The 1st edition of this book was justly acclaimed by teachers and students; within its chosen field it was a peak achievement in exposition. The 2nd edition aims at bringing the work up to date without increasing its size. The result is in many ways a better book than the original. The author says that as it is "essentially elementary" he adheres to the policy of dispensing with full references to the original literature.

The title is important; it both attracts and repels. Many good textbooks embody a personal attitude. Here the crux of the matter is Professor Baldwin's use of "dynamic." What is the content of the word for him? How far does he succeed in communicating that content? The key is to be found on p. 59: "We should do well to recall Hopkins's celebrated aphorism, that *life is a dynamic equilibrium in a polyphasic system.*" The quotation symbolizes for Baldwin his debt to Sir Frederick Gowland Hopkins. When Hopkins realized that the saying had a sparking effect did he decide to leave well alone? We know that he was sensitive to the power of words and the fact that his "aphorism" had fuzzy edges did not perhaps matter much at the time.

Professor Baldwin's book admirably fosters understanding of those aspects of biochemistry which it expounds. It owes much to its readability and occasional racy touches.

The explanation of the roles of energy-rich phosphates faithfully reflects current ideas. Faced with either abandoning Lipmann's very convenient notation for the so-called high-energy bonds or failing to satisfy criticism of the term (and its looser use) Baldwin sticks to $\sim\textcircled{P}$. He is not alone in his dilemma. About a very different matter he writes: "Krebs, however, has grappled with this problem and produced what appears to be a satisfying if not a satisfactory solution." Both the Ayes and the Noes might accept this as describing Baldwin's presentation of bond energy!

The criticisms made here are minor in relation to the major merits of the book as a whole. It is well constructed, coherent and vigorous and the reader's interest is held throughout.

R. A. Morton

Statistical Method in Biological Assay

D. J. Finney. London: Charles Griffin & Co. Ltd., 1952. xix + 661 pages; figures. 22 × 15 cm. £3 8s.

Dr Finney has written an extremely comprehensive, in fact almost a monumental work on biological assay from the statistician's point of view. It is assumed from the outset that the reader has a basic knowledge of statistical method, and this is consequently a book which is aimed primarily at the professional statistician rather than at the amateur who is by training a biologist, in practice a bio-assayist, and of necessity a statistician. Nevertheless, it is to be hoped that this book, which will certainly be a "must" for the professional statistician engaged in the analysis of bio-assays, will reach a much wider public. Unfortunately, although there is much in it of widespread interest, there is also much that will have only a limited appeal, and it is probable therefore that many readers will find it heavy going. The author, recognizing this fact, makes a considerable effort to help readers of different interests by providing several schemes of study, listing chapters of importance in each set of circumstances—for example, his schemes B and C should greatly help the biologist to find his way about this lengthy work.

For many biologists there is a very real philosophical difficulty, to which Dr Finney has obviously given much thought, in the contrast between the mathematical precision of the statistical techniques of design and analysis of bio-assays and the necessarily imprecise knowledge of the substances assayed. The vast majority of assays now carried out are regarded, from the mathematical point of view, as "dilution assays". The basic assumption on which the mathematical theory is based is that the unknown substance that is being assayed against a standard preparation represents a dilution of the standard preparation in an inert material. Many biologists thus tend to feel that the precision of the whole structure of the mathematical theory of bio-assay is built on what they regard as an inadequate foundation; for bio-assay is necessary only for substances that cannot be characterized in terms of weight, i.e., for substances that are not chemically pure. The very chemical impurity of the substance is usually associated with the assayer's inability to say, with any certainty, that the unknown preparation does in fact represent a dilution of the standard preparation in inert material only. The impurities in either the standard or the test preparation may be antagonists or synergists of the active principle that it is desired to measure; moreover, there may be more than one active principle present in one or both substances, and the reaction of different biological systems to these different active principles may vary. This state of uncertainty is, and always must be, present in biological assay since, in removing it, one removes at the same time the need for biological assay. The lack of a precise and firm foundation for an extremely precise mathematical theory is therefore always present in the mind of the biologist: it is not always present in the mind of the mathematician, and consequently one of the very valuable contributions that Dr Finney has made in writing this book is to be found in the way in which he has attempted to reconcile these two points of view.

Thus, although it is true that, occasionally, experimental situations are met which require difficult experimental designs, and the analysis of which can only be done in the rarified atmosphere of the advanced theory of mathematical statistics, the vast majority of assays (as Dr Finney himself points out) are best done with the simplest possible design and analysed by the use of the simplest possible metameters of dose and response that give an adequate linear relationship. The impossibility in most cases of finding any logical reason for choosing these particular metameters is only another argument in favour of simplicity. Dr Finney's discussion of this problem deserves a very wide public indeed.

Many of these difficulties are the direct result of the fact that there is still a gulf, widened indeed by the difficulties of language, between the biologist and the statistician, both of whom are engaged in bio-assay. There is a further result of this dichotomy which Dr Finney also does well to discuss. I refer to the too-ready acceptance by the biologist of statistical precision. Thus, many biologists seem to be so dazzled by estimates of potency which have fiducial limits of error attached to them that, without questioning the ways in which these estimates were arrived at, they accept them as somehow infallible. In chapter XV, Dr Finney discusses once again the whole question of the validity of assays, expanding and enlarging the hypothesis first put forward by Jerne and Wood in 1949. This chapter alone should be available in every laboratory where an assay is carried out and should be read by every biologist engaged in quantitative pharmacology.

One slight omission that may be noted is the absence of any reference to estimates of potency derived from a series of assays, each of which provides from within itself no estimate of error. There are still a very large number of assays, especially in the field of immunology, which depend on methods of this sort. These methods are referred to in the 1953 edition of the *British Pharmacopoeia* under the heading "direct estimates of error". Perhaps it was the very simplicity of the statistics involved which led the author to omit this type of assay, but it is additionally unfortunate that he should use the term "direct assay" to mean what is usually called "assay by graded dose method" or "assay depending on the determination of the effective dose for each animal".

Minor criticisms such as this are, however, unimportant. I have no doubt that Dr Finney's book will find its way on to the bookshelves of every statistician engaged in the design and analysis of bio-assay. It will be invaluable as a book of reference; it is clearly written and it is admirably produced; one must sincerely hope that its very comprehensiveness, that its consequent length, and that, alas, its price, will not prevent it reaching a very wide public.

Walter L. M. Perry

International Review of Cytology. Volume I

Edited by *G. H. Bourne & J. F. Danielli*. New York : Academic Press, Inc., 1952. (Prepared under the auspices of the International Society for Cell Biology.) vi + 368 pages ; illustrations. 23 × 16 cm. £3 2s. 6d.

The widespread interest which is now evident in the study of the localization of enzymes in cells makes it desirable to include in this Bulletin at least a brief outline of this initial volume of a new series, just received for review at the time of going to press. Prepared under the auspices of the International Society for Cell Biology, this and the following volumes of this series are expected to appear annually and are intended to cover the subject of cytology in all its aspects. The plan is to choose a limited number of subjects for consideration in each issue, and to obtain an internationally known authority to deal independently with each of these subjects. Naturally the whole field of cytology will be covered only gradually by this method, but it has the advantage, which is clearly shown in the present volume, of making possible self-contained and reasonably comprehensive reviews of individual aspects of cytological subjects. The drawback to the method is the inevitably disconnected nature of the various chapters; and in fact it is difficult to see any plan in the arrangement of subjects adopted by the editors. The scope, however, is very wide, covering for example essays on nuclear reproduction (C. L. Huskins); nutrition of animal cells in relation to their enzymic capacities (G. W. Kidder); cell enzymes in relationship to membrane permeability (T. Rosenberg & W. Wilbrandt) and to sugar absorption (R. Brown); bacterial cytology (K. A. Bisset) and reproduction of bacteriophage (A. D. Hershey); factors controlling tissue staining (M. Singer) and freeze-drying techniques in cytology (L. G. E. Bell). Finally a very thorough account of histochemistry of esterases is contributed by G. Gomori. There is no doubt that this new annual volume will fill a gap which at present exists in the "annual review" type of publication as far as cytology is concerned.

F. Dickens

* * *

Printed in Great Britain by Spottiswoode, Ballantyne & Co. Ltd., London & Colchester

BRITISH MEDICAL BULLETIN
VOLUME 9 NUMBER 3 1953

		PAGE
Introduction : viruses yesterday, today and tomorrow	*C. H. Andrewes*	169
Ecology of virus diseases	*Sir Macfarlane Burnet*	173
Progress in viral immunology	*Wilson Smith*	176
Chemotherapy of virus diseases	*E. Weston Hurst*	180
Techniques for the study of viruses	*Forrest Fulton*	186
Smallpox and related virus infections in man	*A. W. Downie & A. Macdonald*	191
Recent advances in the study of poliomyelitis, with particular reference to prevention	*A. J. Rhodes*	196
Coxsackie viruses	*J. O'H. Tobin*	201
The common cold	*C. H. Andrewes*	206
The viruses of epidemic influenza	*A. Isaacs*	208
Mumps	*E. T. C. Spooner*	212
Yellow fever : a problem in epidemiology	*G. W. A. Dick*	215
Hepatitis	*F. O. MacCallum*	221
The psittacosis-lymphogranuloma group of viruses	*S. P. Bedson*	226
Q fever in Britain	*M. G. P. Stoker*	231
The post-infection encephalitides	*E. Weston Hurst*	234
Some veterinary diseases of medical interest	*A. W. Gledhill*	237
Book Reviews		242

Scientific Editor: *Forrest Fulton*

Viruses Yesterday, Today and Tomorrow

This number contains some account of the present status of research upon viruses in general; it then deals more particularly with a few viruses which are either of major practical importance or represent "growing points" of scientific advance. The rickettsiae, though not usually classified as viruses, are often considered along with them since they are studied by very similar techniques; it is therefore not out of place that an article[1] is devoted to Q fever, a rickettsial disease now coming to be of practical importance almost everywhere in the world. Psittacosis and related viruses are also dealt with[2], though many people now tend to group these with the rickettsiae, rather than with the "true" viruses: in particular their susceptibility to sulphonamides or antibiotics makes us think of them as belonging in a separate category from the latter.

[1] See Stoker, p. 231.
[2] See Bedson, p. 226.

During the decades following the discoveries of Pasteur and Koch, bacterial causes of one disease after another were described; then a plant disease—tobacco mosaic—and foot-and-mouth disease were found, before the close of the century, to be transmissible agents of infection able to pass through filters that held back bacteria. The modern era of virus research had begun.

In the first twenty years of the present century, advance was not particularly rapid. True, it was recognized that there were diseases, such as smallpox and measles, that had filtrable causes. But there were—very naturally—many attempts merely to extend and develop orthodox bacteriological techniques. People tried to stain viruses, to see them with microscopes of—as we now know—inadequate resolving power, and to grow them in lifeless culture media. Gradually it became clear that most viruses were too small to see with existing optical methods and that they were obligate cell parasites which just would not grow on such things as agar plates. They had to be studied by purely indirect methods, that is, by the effects they produced on the host they attacked. Animal experiment and histology were the virus workers' methods.

At the end of the First World War the need for greatly extended work on viruses became apparent in Britain. The Medical Research Council's National Institute for Medical Research was in its beginnings at Hampstead, London, and it was decided that the study of viruses should be one of its major activities. For several reasons, dog distemper was selected as one of the problems for initial attack; it was hoped that clues would be found which would help to solve the problems of the respiratory infections of man. There resulted Laidlaw and Dunkin's[3] classical work on dog distemper and the means of immunization against it. This led on to the discovery by Smith, Andrewes and Laidlaw of a means of studying human influenza in the laboratory; for the ferret, which had been so useful in the distemper work, was found also to be susceptible to influenza. In fact, as two of our articles show[4 5], human respiratory infections have continued to be major problems under study at the Institute ever since. Elsewhere in London, Bedson[6] was showing that psittacosis, or parrot fever, had a filtrable cause and could be transmitted to mice.

No less important were the new techniques which were developed. Elford's methods of graded filtration were amongst the more important.[7] In the last twenty-five years, yet other methods have come into use one after another. Barnard's development of ultra-violet photomicrography, at the National Institute for Medical Research, Hampstead, London, was followed by the invention of the electron microscope, now at last coming to be used to solve definite problems and not merely to produce pretty pictures. Many viruses were found to grow in one way or another in fertile hen-eggs. Influenza viruses, growing in the allantoic cavity, were found to become adsorbed to and to agglutinate red blood cells of birds and mammals. Now, by ingenious modifications of technique, more and more viruses are being shown to be capable of agglutinating red cells. These findings have opened up new possibilities of studying the reaction between a virus and its host by means not involving costly and cumbersome animal experiment. Similar hopes follow from more recent work on growth of viruses in tissue cultures. Poliomyelitis[8], and some other viruses, can be titrated and serologically typed by methods based on their power to produce death and destruction of mammalian cells growing in vitro.

Development of new laboratory techniques has not meant that older, cruder methods can always be discarded. Where convenient laboratory animals are not susceptible, it may be still necessary to test for presence of virus by observing the effects of deliberate inoculation of human volunteers.[5 9] Such methods, apparently simple yet presenting their own difficulties, may have to be carried out *pari passu* with newer laboratory techniques. Study of the epidemiology of virus diseases may demand very refined serological methods[4] or the co-operation of skilled entomologists.[10]

Recent investigations into the infection of bacterial cells by bacterial viruses or bacteriophages may prove to have wide implications for the virus diseases of animals. We have lately learnt much about how the phages attach themselves to bacteria, how only a part of the phage penetrates to the interior of the attacked cell, its "overcoat" being left outside, how its infectivity for other cells temporarily disappears but is regained when a complete virus is once more reconstituted. Neither simple binary fission nor synthesis on a template will fit the facts as to its mode of multiplication. It now seems that two similar viruses infecting one cell may interchange genetic elements, so that a sort of hybrid virus results. Virus genetics seems to be a subject for the future and new viruses may even be contrived and "tailored to measure". Biochemistry is being increasingly called in to help unravel the mysteries of how viruses operate. Less is heard now than a few years ago as to what viruses are and where they come from: such speculation is felt to be rather idle while there are so many new and exciting things to discover about how they produce their effects. Such deeper understanding seems sure to lead, amongst

[3] See Gledhill, p. 237.
[4] See Isaacs, p. 208.
[5] See Andrewes, p. 206.
[6] See Bedson, p. 226.
[7] See Fulton, p. 186.
[8] See Rhodes, p. 196.
[9] See MacCallum, p. 221.
[10] See Dick, p. 215.

other things, to a rational approach to the chemotherapy of virus diseases. At present, in this field, there are hints that we may one day be able to attack even the small viruses[11], but there is nothing for us yet of a practical nature. It is our present helplessness to cure most virus diseases of animals which makes these diseases increasingly important. With bacterial and protozoal diseases coming under better and better control, the viruses loom relatively larger in their importance for man's health and happiness. It is clear that workers trained in very many different disciplines will be needed to solve the many knotty problems they present.

C. H. Andrewes

[11] See Hurst, p. 180.

We acknowledge our indebtedness to Dr C. H. Andrewes for his help and advice in the planning of this symposium.

Notes on Contributors

DR C. H. ANDREWES is now Deputy Director of the National Institute for Medical Research at Mill Hill, London. He has previously contributed two articles to *British Medical Bulletin*: the first, with which a biographical note was published, was entitled "Virus diseases of man: a review of recent progress" (*Brit. med. Bull.* 1944, **2**, 265), and the second, which Dr Andrewes wrote in collaboration with Dr Harold King, was on the chemotherapy of rickettsial and virus diseases (*Brit. med. Bull.* 1946, **4**, 272).

PROFESSOR S. P. BEDSON obtained his scientific and medical education at the University of Durham and at the Institut Pasteur, Paris. Since 1949 he has been part-time Adviser in Pathology to the Ministry of Health, and he has recently also taken up part-time research work in the Bland-Sutton Institute of Pathology at the Middlesex Hospital, London, where the study of viruses is being continued. He was Assistant Bacteriologist at the Lister Institute of Preventive Medicine, London, from 1921 to 1926; for the first three years of this period he studied the blood platelet and experimental purpura, and for the remaining two years was seconded for work with the Foot-and-Mouth Diseases Committee. He was Freedom Research Fellow at the London Hospital from 1926 to 1934, after which he became Professor of Bacteriology in the University of London at the London Hospital and Medical College; he held this appointment until 1952. During the years he spent at the London Hospital and Medical College, Professor Bedson worked with several viruses, including those of herpes and psittacosis. He was elected into the Fellowship of the Royal Society in 1935 and was awarded the Conway Evans Prize in 1952. He is joint author with A. W. Downie, F. O. MacCallum & C. H. Stuart-Harris of *Virus and rickettsial diseases* (1950), and has published numerous scientific papers on the blood platelet and purpura and on viruses.

SIR MACFARLANE BURNET was, from 1928 to 1943, Assistant Director of the Walter and Eliza Hall Institute of Medical Research, Royal Melbourne Hospital, and since then has been Director of the Institute. He is also Professor of Experimental Medicine in the University of Melbourne. From 1926 to 1927 he worked at the Lister Institute of Preventive Medicine, London, and from 1932 to 1933 at the National Institute for Medical Research, London. Sir Macfarlane Burnet's work has been concerned mainly with bacterial viruses and, since 1932, with animal viruses, particularly those of influenza. At present he is interested chiefly in the study of the intrinsic characteristics of influenza viruses. He was elected into the Fellowship of the Royal Society in 1942 and was awarded the Royal Medal in 1947; three years later he delivered the Croonian Lecture. His main publications are *Natural history of infectious disease* (2nd ed., 1953), of which a second edition is in the course of publication, and *Virus as organism* (1945).

DR G. W. A. DICK was educated at Edinburgh University and at the Johns Hopkins University. Before the Second World War he was Assistant Pathologist at Edinburgh Royal Infirmary. From 1940 to 1946 he was in the Royal Army Medical Corps as a specialist in pathology. From 1946 to 1951 he was in the Colonial Service at the Virus Research Institute, Entebbe, Uganda. He was for some time a Rockefeller Foundation Scholar and Research Fellow in epidemiology at the Johns Hopkins University. His investigations have been mainly concerned with yellow fever, poliomyelitis and some of the recently isolated African viruses; he has published papers on all of these subjects. In 1949 Dr Dick was awarded a gold medal by Edinburgh University for his M.D. thesis on the Mengo group of viruses. He is a specialist consultant to the Communicable Diseases Centre, Atlanta, Georgia, USA, and at present is on the staff of the Medical Research Council at the National Institute for Medical Research, London, where he is engaged on the study of virus hepatitis.

PROFESSOR A. W. DOWNIE is at present Professor of Bacteriology at the University of Liverpool, where he is working with pox viruses. He was assistant to the Professor of Bacteriology at the University of Aberdeen from 1923 to 1927, when he became Lecturer in Bacteriology at the University of Manchester; he held this post for seven years and then, from 1934 to 1939, was Senior Freedom Research Fellow at the London Hospital. For the first year of this period he was Voluntary Assistant at the Rockefeller Institute for Medical Research, New York. From 1939 to 1943 he was Senior Member of the Scientific Staff at the National Institute for Medical Research, London. Approximately half of the 50–60 scientific papers written by Professor Downie deal with various bacteriological topics, particularly infection and immunity in bacterial infections; the remainder of the papers are mostly concerned with virus research, particularly in relation to the pox diseases. He is joint author with S. P. Bedson, F. O. MacCallum & C. H. Stuart-Harris of *Virus and rickettsial diseases of man* (1950), and author of "Experimental streptococcal infection and immunity" (*J. Path. Bact.* 1930, **33**, 563) and of "Infection and immunity in smallpox" (*Lancet*, 1951, **1**, 293).

DR FORREST FULTON is Reader in Bacteriology and Immunology in the University of London at the London School of Hygiene and Tropical Medicine. From 1939 to 1942 he was in the Emergency Public Health Laboratory Service at Oxford, and from 1942 to 1948 was a member of the Scientific Staff of the Medical Research Council at the National Institute for Medical Research, London. His present work is concerned wholly with viruses, but he has published a number of papers on bacteriological problems and on rickettsiae. These include "Staphylococcal enterotoxin: with special reference to the kitten test" (*Brit. J. exp. Path.* 1943, **24**, 65); in collaboration with A. M. Begg, "The antigenic structure of typhus rickettsiae" (*Spec. Rep. Ser. med. Res. Coun., Lond.* No. 255, p. 163); and with L. Joyner, "The cultivation of *Rickettsia tsutsugamushi* in the lungs of rodents: preparation of a scrub-typhus vaccine" (*Lancet*, 1945, **2**, 729).

DR A. W. GLEDHILL, after spending three years in Uganda as a veterinary officer concerned chiefly with the control of epizootic diseases, held a research post for eight years at the Institute of Animal Pathology, University of Cambridge. He joined the staff of the National Institute for Medical Research, London, in 1947. Since leaving Africa in 1938, he has been concerned principally with investigating swine and human influenza, the virus diseases of cats and the antigenic structure of *Erysipelothrix rhusiopathiae*. At present he is engaged with colleagues in investigating the properties of a virus causing hepatitis in mice, and in the attempted adaptation of the viruses of human hepatitis to laboratory animals. He has published papers on these and other subjects.

DR E. WESTON HURST was educated at King Edward's School, Birmingham and at Birmingham University. From 1923 to 1926 he did external postgraduate work as Walter Myers Travelling Student of the University of Birmingham and for the Medical Research Council. He was a member of the research staff of the Lister Institute of Preventive Medicine, London, from 1928 to 1936; and of the Rockefeller Institute, Princeton, from 1932 to 1934. He was Reader in Experimental Pathology at the University of London from 1932 to 1936. From 1936 to 1943 Dr Weston Hurst was Director of the Institute of Medical and Veterinary Science, Adelaide, during which time he also became Keith Sheridan Professor of Experimental Medicine at the University of Adelaide. Since 1943 he has been Head of the Sub-Division of Micro-Organismal Research in the Biological Laboratories of Imperial Chemical (Pharmaceuticals) Ltd., where

NOTES ON CONTRIBUTORS

he is engaged in research on the chemotherapy of virus diseases. In the past his research work was concerned with neuropathology and virus diseases. Dr Hurst was William Withering Lecturer to the University of Birmingham in 1935, Sir Joseph Bancroft Orator to the Queensland Branch of the British Medical Association in 1940 and the first G. E. Rennie Lecturer, in 1941, to the Royal Australasian College of Physicians. His publications include papers in various scientific and medical journals.

DR A. ISAACS has been, since 1950, on the Scientific Staff of the Medical Research Council, working at the National Institute for Medical Research, London. He has previously held a McCunn Research Studentship at the Department of Bacteriology, Glasgow University, a Medical Research Council Studentship at the Department of Medicine, University of Sheffield, and a Rockefeller Travelling Fellowship at the Walter and Eliza Hall Institute of Medical Research, Melbourne. His published work has been mainly concerned with theoretical aspects of influenza virus growth and with the epidemiology of influenza. He is at present responsible for laboratory investigations carried out at the World Influenza Centre as part of the World Health Organization programme on influenza.

DR F. O. MACCALLUM has previously contributed to the Bulletin articles on the use and function of the Virus Reference Laboratory of the Public Health Laboratory Service (*Brit. med. Bull.* 1951, **7**, 174) and on infective hepatitis (*Brit. med. Bull.* 1943, **1**, 112), with which biographical notes were published. He has been Director of the Virus Reference Laboratory since 1946 and since then has reported work on lymphocytic choriomeningitis, poliomyelitis, Q fever and smallpox. He was one of the authors of *Infective hepatitis : studies in East Anglia during the period 1943–47* (*Spec. Rep. Ser. Med. Res. Coun., Lond.* No. 273), and a member of the World Health Organization Expert Committee on Hepatitis.

DR A. MACDONALD is Reader in Bacteriology at the University of Liverpool, where he is investigating the factors in normal mammalian sera that are capable of modifying experimental virus infections. From 1938 to 1942 he was Garden Research Fellow, University of Aberdeen, and was Director of the Public Health Laboratory Service at Norwich from 1944 to 1948, when he took up his present post. He was previously concerned with investigations on a number of bacterial infections, particularly those of public health importance, and on toxoplasmosis and the pox group of viruses. His publications include a paper on the serological diagnosis of human toxoplasmosis (*Lancet*, 1949, **1**, 950) and he is joint author with A. W. Downie of "A study of the pox viruses by complement fixation and inhibition of complement-fixation methods" (*J. Path. Bact.* 1950, **52**, 389) and with G. M. Lloyd & R. E. Glover of an article on human infection with the virus of contagious pustular dermatitis (*Lancet*, 1951, **1**, 720).

PROFESSOR A. J. RHODES was, from 1934 to 1941, Lecturer in Bacteriology at Edinburgh University and, from 1945 to 1947, held a similar post at the London School of Hygiene and Tropical Medicine. At present he is Director of Research at the Hospital for Sick Children, Toronto; Professor of Virus Infections, School of Hygiene, University of Toronto; and Honorary Consultant, Connaught Medical Research Laboratories, University of Toronto. Professor Rhodes is joint author, with C. E. van Rooyen, of *Virus diseases of man* (2nd ed., 1948) and of *Textbook of virology* (2nd ed., 1953); he has also published numerous articles on virus infections, particularly on poliomyelitis and Coxsackie infections.

PROFESSOR WILSON SMITH is at present Professor of Bacteriology at University College Hospital Medical School, University of London. He was educated at the University of Manchester and, after some experience of general practice and a postgraduate course in bacteriology, began his research career at the National Institute for Medical Research, London. There, over a period of 12 years, he worked on various problems in both bacteriology and virology. Work in collaboration with the late Sir Patrick Laidlaw and Dr C. H. Andrewes resulted in the discovery of the influenza virus. This discovery made possible the intensive influenza research of the last 20 years, in which Professor Wilson Smith has continued to take an active part. In 1939 he was appointed to the Chair of Bacteriology at the University of Sheffield, which he occupied for the duration of the war. In spite of heavy commitments connected with the war effort, it was in this period that his researches into the mechanisms of staphylococcal infection and the role of staphylocoagulase were carried out. In 1946 he accepted an invitation to the Chair that he now occupies and he is at present engaged in investigations on the constitution and structure of the influenza viruses. In 1949 he was elected a Fellow of the Royal Society.

PROFESSOR E. T. C. SPOONER has, since 1947, been Professor of Bacteriology and Immunology at the London School of Hygiene and Tropical Medicine. He was educated at Clare College, Cambridge and at St Bartholomew's Hospital, London, where, from 1927 to 1928, he was House Physician. From 1929 to 1931 he was Commonwealth Fellow at Harvard Medical School. In 1929 he became a Fellow of Clare College, Cambridge, and, in 1939, Tutor of Clare College; he held these appointments until 1947, when he took up his present position. From 1931 to 1946 he was University Demonstrator and Lecturer in the Department of Pathology at the University of Cambridge. For a year (1942–43) during the Second World War, he held a temporary commission in the Royal Army Medical Corps, and then spent the following year as Director of the Emergency Public Health Laboratory at Cambridge. He has published papers on tetanus, certain virus diseases and wound infection.

DR M. G. P. STOKER was, in 1946, attached as Pathologist to General Headquarters (India) Medical Research Organization, on the staff of the Base Typhus Research Unit at Poona. In 1947 he became University Demonstrator in Pathology at the University of Cambridge, and a year later was appointed University Lecturer in Pathology and was elected a Fellow of Clare College. In 1949 he became Director of Medical Studies and Assistant Tutor at Clare College. Since 1951 he has been Huddersfield Lecturer in Special Pathology at Cambridge. In the latter part of the Second World War he worked on the rickettsial diseases of India with Dr D. R. Seaton; they were particularly interested in Indian tick typhus and, by using specific serological methods, showed a relationship to Rocky Mountain spotted fever. Since going to Cambridge, his work has been chiefly with Dr R. R. A. Coombs on serological methods in virus systems and, with Drs J. A. R. Miles and D. Surrey Dane, on a virus epizootic in Manx shearwaters. During the last four years, Dr Stoker's work has been on Q fever, chiefly in association with Dr B. P. Marmion. This work has included the epidemiological study of the disease in Britain, as well as the laboratory investigation of the causative agent. His published work includes papers on serological analysis of typhus cases in India, application of the conglutinating complement-absorption test to virus systems, and Q fever in Great Britain.

DR J. Ó'H. TOBIN studied medicine at the University of Oxford and at the Radcliffe Infirmary, Oxford. After serving in the Royal Army Medical Corps during the Second World War, he took the Diploma in Bacteriology in 1948 at the University of Manchester, since when he has been a Lecturer in Bacteriology there. He has investigated virus diseases occurring in the north-west of England, and has published a number of papers on this subject.

ECOLOGY OF VIRUS DISEASES

SIR MACFARLANE BURNET F.R.S.

The Walter and Eliza Hall Institute of Medical Research, Melbourne, Australia

1 Basic features of viruses
2 Patterns of virus survival
  a A short-lasting infection with a period of only a few days of infectivity for others, and followed by virtually lifelong immunity
  b Superficial infections in which generalization is not an essential part of the survival cycle
  c Low-grade persisting infection
3 Transfer from one host to another
References

It is in many ways only an affectation to assume that there is a modern ecological approach to the problems of infectious disease which differs essentially from classical epidemiological methods. Yellow fever was being soundly dealt with years before the word ecology began to have common currency in medical and public health circles. Nevertheless it does seem worth while from many points of view to apply consciously to infectious disease, whether due to viruses or other pathogens, the concepts that have been developed in regard to the behaviour, distribution and population densities of living organisms. Ecology is concerned with any organism, not as an individual to be analysed, but as a continuing population of individuals, community or species, in its relation to its environment. The major concern of the ecologist with fluctuations in numbers and geographical distribution finds an obvious parallel in the varying intensity and movement of disease.

Modern evolutionary theory is based on a combined genetic and ecological approach. Evolution amongst micro-organisms can have a time-scale so compressed that highly important changes may take place within a few years. The serological changes in influenza virus A since 1933 can be instanced. Ecological discussion of virus disease must necessarily therefore take account of the genetic mutability of the organisms concerned. From a rather academic view-point the humanly important manifestations of virus activity—epidemics, prevalence or individual cases of disease—need for their complete understanding a description of the associated changes in the populations of virus particles. Any description of a biological population, in addition to information on its density and spatial distribution, must also include an account of the distribution of any significant genetic sub-types. In virus work this approach is of no importance at the practical level of handling a case or an epidemic of a given disease but it may be very significant indeed in helping toward an understanding of the persistence of virus diseases and changes in them over long periods of time.

Looked at from this angle of population changes, we have two different systems to consider. When an infection is initiated in a susceptible host we have something comparable to the entry of a species into a limited environment previously free of its kind. The degree to which the virus will multiply, its effect on the host—whether it will produce death or allow recovery—and in the event of recovery the time course of virus increase and decline, and the period for which viable virus will remain in the body—all these factors need to be understood. It is orthodox to consider this in terms of tissue susceptibilities, transport of virus in the body and the development of immunity. Nevertheless it is basically an ecological problem—the influence of a given environment on the population history of a micro-organism. For the present purpose, however, we can consider the result as depending essentially on three factors: the genetic nature of the virus (i.e., its intrinsic virulence), the genetic constitution of the host individual (i.e., its intrinsic susceptibility) and the past experience of the host with immunologically similar infective agents.

The second system is concerned not with what happens in individual host animals but essentially with the way some individuals of the parasite species are enabled to survive in between the population peaks of a clinical attack or an epidemic. It is this approach which may often lead to the conclusion that the clinical human disease which is our primary interest has no significant part whatever in the survival of the virus species. Most of the aspects of virus disease that are of ecological interest fall within this second system and only incidental mention need be made of the complex struggle for existence within the individual host.

1. Basic Features of Viruses

The important qualities of human pathogenic viruses as viruses are:

i. The fact that they can multiply only within the living cells of a susceptible host. There are many which in nature are limited to human cells only, others which may multiply in the cells of a variably extensive range of other warm-blooded vertebrates or of arthropod vectors.

ii. All viruses that have been adequately studied have been shown to be subject to inheritable variation; some, like the influenza viruses, are highly mutable. There is increasing evidence (not, however, sufficient to convince all virologists) that mutations are essentially spontaneous and that their apparently adaptive character is wholly a result of the selective survival of appropriate mutants.

iii. As is commonly the case with specialized parasites, only an almost infinitesimally small proportion of the individuals produced find opportunity to continue the species. An influenza epidemic at its height might mean the temporary existence of 10^{20} viable influenza virus particles in a large city. A few weeks later the only viable descendants of that multitude might well be the insignificant number stored in dry ice in the local virus laboratory. The thread by which one influenza epidemic follows another may often be as tenuous as the transfer of only one virus particle from one individual to another. This regeneration of populations after catastrophic reduction in numbers has often been discussed by ecologists and population geneticists concerned with higher organisms.

iv. With some important exceptions, most viruses are readily destroyed by environmental agencies such as exposure to light and desiccation. For transfer of virus from one host to another it is therefore necessary that material containing virus—blood excretions or discharges—should be conveyed rather directly to the new host in such a way that virus can reach and infect the susceptible cells.

v. For many viruses a given host individual is available only once in its lifetime and for a short time only. Acquired immunity in the host is an important limitation to virus activity and is the dominant factor in determining the epidemiology of many virus diseases.

2. Patterns of Virus Survival

Within the limitations imposed by these qualities there are few patterns of possible survival that are not exploited by one or other virus. If we confine ourselves to human virus diseases, three major patterns can be recognized.

a. A Short-Lasting Infection with a Period of only a Few Days of Infectivity for Others, and Followed by Virtually Lifelong Immunity

This is perhaps the standard form of virus disease. If the circumstances are such that human infections can readily occur, such diseases are characteristically limited to young people. It is an obvious deduction amply verified by experience that, the more apt conditions are for the spread of infection, the younger will be the characteristic age of infection and immunization.

There are two main divisions of this group:

i. *Virus diseases transferred by blood-sucking vectors.* In dengue and urban yellow fever the only vertebrate involved is man: infection of *Aedes* mosquitoes is from human patients in the short-lasting viraemic phase. The mosquito-borne encephalitides are rather closely related to dengue and yellow fever but in the mosquito-borne encephalitides the main part of the cycle of survival is in non-human hosts. Birds are the important type of natural vertebrate host and in them infection in most instances appears to be symptomless but is associated with a sharp transient viraemia and is followed by long-lasting immunity.

ii. *The virus exanthemata.* These are infections which spread readily by salivary contamination or droplet inhalation and are associated with a regular secondary phase of generalization by the blood stream. Although the nature of the infection at initiation is unknown, one must postulate, for all of these infections, an initial phase in which cells in contact with the environment are infected. Despite the suggestive analogies in Fenner's study of the pathogenesis of mouse-pox (see Fenner, 1948), there are still large gaps in our understanding of what takes place between initiation of infection and the culminating outburst of symptoms, exanthem and infectivity for others.

The common features of these superficially rather dissimilar types of disease are (i) the existence of a viraemic phase and, perhaps as a result, a long-lasting immunity, (ii) a short-lasting period of infectivity and (iii) the elimination of the virus from the body after recovery. In addition to providing protection against symptomatic infection, immunity in this group is sufficient to ensure that an immune individual will not become subclinically infected, and will not therefore be able to serve as a link in a chain of infection. There is still some controversy as to whether virus is in fact wholly eliminated after measles and such other infections as give many years' immunity. Some regard it as conceivable or likely that virus remains present, multiplying slowly, and providing sufficient antigenic stimulus to maintain a high level of antibody. The argument against this is in part at a purely immunological level which does not concern us here. At the epidemiological level there are even stronger arguments. Virus can act as an immunological stimulus only when it escapes from the cells in which it multiplies. This should also provide sooner or later an escape of virus into the environment, with the initiation of infection in susceptibles. The history of measles in isolated communities such as the Faröe Islands provides a categorical negative to this potentiality. Those who have been infected in the past remain immune, the new susceptibles never suffer from measles nor lose their susceptibility until measles is again introduced into the community from without.

Survival of the virus in diseases of this " typical " character depends essentially on the efficiency of the transfer process. The donor of virus is available as such only for a few days. Susceptible hosts are automatically eliminated as they suffer an attack, and in general the susceptible population is composed of young individuals. At a certain point in any epidemic the remaining population of susceptibles is too low to allow spread to continue, and unless the effective size of the epidemiological unit is beyond a certain magnitude the disease will disappear until it is re-imported from some other epidemiological unit.

A very similar situation may arise in regard to disease in which the normal host is some vertebrate other than man. The example most familiar to us in Australia is Murray Valley encephalitis, a mosquito-borne disease that in the last 40 years has made its appearance only four times in the Murray Valley of southern Australia (Anderson, 1952). Here the requirements for spread are the presence of birds in the short-lasting viraemic phase of infection, mosquitoes of appropriate species and at an environmental temperature that will allow virus to complete its biological cycle in the vector, and non-immune birds or animals to which the virus can be passed. While such movement is actively proceeding human cases can occur.

In an epidemic season like that of January to April 1951 the virus flourishes, leaving evidence of infection in a large proportion of the human mammalian and avian hosts in the affected region. In the winter the cycle for one reason or another is interrupted, and, from its very nature, once interrupted it cannot begin again unless and until the virus can be re-imported from a region in which no such interruption occurs. The failure of Anderson and his colleagues to find any evidence of Murray Valley encephalitis virus in early 1952, in the area where it flourished 12 months earlier, provides very strong evidence of the importance of this point of view. Summer conditions in the area are such that the virus can spread with the greatest of ease—if birds infected one year continued to carry the virus indefinitely or if a secondary over-wintering vector like the mite existed, then yearly outbreaks must have been inevitable.

b. Superficial Infections in which Generalization is not an Essential Part of the Survival Cycle

Of the human virus diseases that have been adequately studied, influenza and poliomyelitis fall into this group, but in all probability there are a number of other virus infections of respiratory or alimentary tracts which behave similarly. For the present the influenza viruses must be regarded as having no significant host other than man. The possibility that the work of Shope (1943) on the earthworm-lungworm cycle of swine influenza will sooner or later lead to the recognition of an analogous primary cycle for the human influenza viruses cannot be excluded. Current knowledge of the human disease does not, however, appear to call for any such hypothesis.

Influenza is an infection readily transferred by droplet transfer to fresh susceptibles. It differs essentially from the exanthemata in having no significant capacity to multiply elsewhere than in the respiratory epithelium, and as a result little virus passes into the blood. Antibody production, probably for this reason, is less long-lasting. Further, the effectiveness or otherwise of the antibody will be determined by the concentration in which it is present or by the film of moisture lying over the cells lining the respiratory tract. Under these circumstances it becomes a " practical proposition " for the virus to make a second or subsequent visit to the same host. Influenza virus A (or B) characteristically survives by passage through human communities partially immune from past experience of similar influenza. If a strain of virus arises which has changed its serological type significantly from that of the strain responsible for the preceding visitation, this will be at an advantage in overcoming the obstacle of persisting partial immunity. The history of serological change in influenza A since 1933 as recently demonstrated by Hirst (1952) is one of a persistent evolution of serological novelty. Since 1933 there have been seven major sub-types recognized: WS, PR8, MEL, TAL, ALA-41, FM1 and the 1951A prime strains. With only occasional exceptions each type has had its year or two of dominance over the whole world, and then has been universally replaced by the new type.

Poliomyelitis[1] is not as well studied as influenza but we shall probably be not far wrong in thinking of it as something intermediate between the exanthemata and influenza. Basically poliomyelitis is an infection of the lining of the bowel, prone to pass more deeply into the body, in part carried by the blood stream, in part by passage along nerve paths. The result of extensive serological surveys like those of Paul, Melnick & Riordan (1952) has left the picture a little dubious. Some of the evidence suggests that a single infection will provide immunity lasting for 20 to 40 years or more, but other findings make it more likely that questions of strain, dosage and repetition of infection all play their part in determining the immune status of the individual. With the recent development of much more powerful technical methods for the study of the poliomyelitis viruses and the corresponding antibodies, a much clearer picture should soon emerge. It is highly probable that the Coxsackie group[2] of intestinal viruses will be found to have an ecology very similar to that of the poliomyelitis viruses.

c. Low-Grade Persisting Infection

There are a few viruses which have acquired the survival pattern of transfer from one generation to the next. In the case of man this demands a persistence of infection for 20 or 30 years and herpes simplex is the only available example. The existence of Brill's disease, however, indicates how a similar mechanism could conceivably play a part in the survival of the typhus rickettsia. According to the interpretation of herpes simplex given by Anderson & Hamilton (1949), the infant is normally born from a mother with circulating antibody. It is immune for 6 to 12 months and will usually become infected either with clinical stomatitis or subclinically between the ages of 9 and 24 months. In a normal family, infection will usually be derived from the mother at a time when an outbreak of herpes on her lips coincides with opportunity for salivary infection of some minor trauma in the infant's mouth. After recovery from primary herpetic infection, skin cells in regions involved in or adjacent to the primary lesions carry a latent infection perhaps essentially analogous to the lysogenic state in bacteria. Just as irradiation with ultra-violet light will provoke lysogenic bacteria to lysis, so a variety of stimuli will activate herpes and lead to shedding of infective virus.

This characteristic mode of transfer makes herpes simplex almost the only virus that we can safely regard as having evolved with man throughout his whole history. The fact that a related virus (B virus of Sabin & Wright, 1934) is almost equally common amongst rhesus and cynomolgus monkeys has an interesting bearing on the evolution both of man and of the herpes virus.

Two occasional human pathogens seem in their normal hosts to have developed a similar transfer from one generation to the next. Lymphocytic choriomeningitis virus was found by Traub (1939) to be transmitted to young mice in utero in certain mouse colonies, and to persist throughout life in the tissues of the animal.

Psittacosis either in parrots in the wild or in budgerigar-breeding aviaries is usually contracted in the nest, presumably because of a temporary activation of a latent infection in the laying female. There is a tendency for all members of the psittacosis group of viruses, including trachoma and lymphogranuloma venereum, to produce low-grade persisting infections with poor immunological response and prolonged potential infectivity for others.

3. Transfer from One Host to Another

For every virus we can attempt to describe the way in which the species survives in nature. In general the more effective our attempt is, the more readily can we control the disease concerned. When our ignorance is as complete as it is about the natural history of the virus of homologous serum jaundice, it is usual to find a similarly complete impotence in handling the disease.

A large proportion of the insect-transmitted viruses which can cause disease in man survives by passage through a cycle which normally is confined to non-human mammalian or avian hosts. Human infection is an accident which plays no part in the chain of survival. Urban yellow fever[3] almost certainly depends on the existence of the jungle cycle for its persistence from epidemic to epidemic, although during the currency of an epidemic the human-mosquito cycle only is significant.

Transfer of a virus from one vertebrate host to a completely distinct species provides some very important problems in the fields of medicine, veterinary medicine and wild-life control. Generalizations are difficult to make. Probably the most generally useful is that where a virus has established a stabilized ecological relationship with whatever hosts and vectors are concerned in its survival, the effect of the virus on infected individuals is usually trivial. Major epidemics or epizootics are always an indication of an abnormal situation. When a thoroughly stabilized virus is transferred to a new host the outcome is impossible to predict by any a priori rules. It is probably the commonest result to find that no detectable effect whatever follows infection. The virologist's difficulty in finding appropriate species in which to study human

[1] See Rhodes, p. 196 of this Bulletin.
[2] See Tobin, p. 201 of this Bulletin.
[3] See Dick, p. 215 of this Bulletin.

Continued at foot of page 176

PROGRESS IN VIRAL IMMUNOLOGY

WILSON SMITH M.D. F.R.S.

Department of Bacteriology
University College Hospital Medical School, London

1 Development of serological diagnostic techniques
2 Virus haemagglutination
3 Antigenic complexity of viruses
4 Antigenic lability
5 The interference phenomenon
References

Immunological research in virus diseases has from the outset been chiefly concerned with two objectives: (i) the elucidation of the mechanisms of antiviral immunity as a means of unravelling the mysteries of the host-virus relationship, and (ii) the development of practical methods of diagnosis, prophylaxis and specific therapy. The general climate of opinion on the nature of immunity to virus infections has passed through various phases. In the early days of systematic virus research there was a tendency to overemphasize the differences between viral and bacterial infections. It was frequently asserted, though on quite insufficient evidence, that with few exceptions recovery from virus diseases resulted in solid lifelong immunity[1], and it was suggested that this immunity must be based upon mechanisms fundamentally different from those underlying the acquired resistance to bacterial diseases and especially upon cellular as distinct from humoral changes. Protective and virus-neutralizing serum antibodies were readily demonstrable but attempts to apply the various in-vitro serum reactions that had been so successfully exploited in bacteriology were often fruitless and in many cases led to the publication of directly contradictory results. This, coupled with the recognition that virus neutralization was not dependent upon any firm union between virus and neutralizing antibody, led to grave doubt that viruses were capable of eliciting the production of antibodies similar in nature to those associated with bacterial immunity. Failure to confer satisfactory degrees of resistance on experimental animals by means of killed virus vaccines, and the difficulty of conferring effective passive immunity, tended to reinforce this idea of fundamental difference between viral and bacterial immunity.

This early phase was followed by a period of intensive experimentation, mainly by workers in Britain and in the USA, during which the supposed differences were shown to be largely fictitious. The conception of a cellular immunity not dependent upon, or associated with, production of circulating antibodies was vigorously upheld by Levaditi and his collaborators (Levaditi, 1926, 1946), but most virologists accepted the rapid accumulation of experimental facts which pointed to the conclusions that immunity in virus diseases is essentially of the same nature as immunity in bacterial infections, that in both cases acquired specific resistance is associated with and largely dependent upon the production of circulating antibodies, and that antiviral and antibacterial immune bodies are essentially of the same nature, the apparent differences between them being due to special properties of viruses such as their small size and obligate intracellular parasitism. This changed outlook led to renewed efforts to apply the classical techniques of serology to practical problems of virology and each success in this direction tended to reinforce the idea that viral immunology had nothing new to offer.

Certain recent discoveries, however, have rendered this position untenable, so that current opinion represents at least a partial reversion to the early conceptions. Most far-reaching in their effects have been the discoveries of virus haemagglutination (Hirst, 1941; McClelland & Hare, 1941) and of the phenomenon of virus interference (Hoskins, 1935;

[1] See Burnet, p. 173 of this Bulletin.

ECOLOGY OF VIRUS DISEASES

Continued from page 175

infections is sufficient indication of this. It is quite common, where the natural infection has a variety of alternative hosts, for the human species to react with only a transient viraemia —the usual situation in most of the mosquito-borne encephalitides. The most important case is where the trivial disease of the natural host gives rise to a highly fatal one in a new host usually rather closely related to the natural one. Three examples that may be mentioned are: the fatal human infections due to virus B of monkeys, African hog cholera derived from a mild wart-hog disease, and myxomatosis of the European rabbit, derived from the trivial disease of the South American *Sylvilagus*. The last example is of particular interest at the present time in Australia. Since the virus of myxomatosis was effectively liberated in December 1950 (see Lines, 1951), it has spread widely over the south-east and a current estimate is that the rabbit population of Victoria, Australia, has been reduced by about 70%.

The changes already evident in the increasing proportion of surviving immune rabbits provide a problem of extreme interest to the ecologist. This aspect is now being closely studied and the process by which an equilibrium of one sort or another will be reached can hardly fail to throw light on general epidemiological theory. Irrespective of whether the object is to prevent or, as in this instance, to favour the spread of infection, the primary requirement is an adequate understanding of the ecological situations involved.

REFERENCES

Anderson, S. G. (1952) *Med. J. Aust.* **1,** 97
Anderson, S. G. & Hamilton, J. (1949) *Med. J. Aust.* **1,** 308
Fenner, F. (1948) *Brit. J. exp. Path.* **29,** 69
Hirst, G. K. (1952) *J. exp. Med.* **96,** 589
Lines, E. W. (1951) Reported in: " Medical Societies ", *Med. J. Aust.* **1,** 813
Paul, J. R., Melnick, J. L. & Riordan, J. T. (1952) *Amer. J. Hyg.* **56,** 232
Sabin, A. B. & Wright, A. M. (1934) *J. exp. Med.* **59,** 115
Shope, R. E. (1943) Messenger Lecture in : Rivers *et al.* *Virus diseases*, p. 85. Cornell University Press, New York
Traub, E. (1939) *J. exp. Med.* **69,** 801

Findlay & MacCallum, 1937). These have led to a restatement of the concept of cellular immunity as a factor in acquired host resistance which may operate independently of the humoral factors.

1. Development of Serological Diagnostic Techniques

Until satisfactory methods of virus purification and concentration were evolved, virus neutralization remained the only demonstrable serological reaction in most of the virus diseases of man and animals. It still remains the most dependable method for assessing the level of immunity enjoyed by the serum donor but has the great disadvantage that it requires the use of susceptible experimental animals. Its scope was greatly extended by the introduction of the embryonated egg into virus research and particularly by the development of numerous alternative egg techniques for virus cultivation and study (Woodruff & Goodpasture, 1931; Burnet, 1936; Beveridge & Burnet, 1946). Virus-neutralization tests in eggs in many cases introduced a new order of precision into antibody assay; for example, the accuracy of the pock-count method with viruses like vaccinia, which produce focal lesions on the chorio-allantoic membrane[2], is limited only by the number of eggs available for the test.

Early claims that in-vitro serum reactions like agglutination and complement fixation could be demonstrated with viral antigenic materials and specific antisera were met with scepticism, but today they are acknowledged to have been fully justified. Indeed these in-vitro reactions now play a very important and an ever-increasing role in the laboratory diagnosis of virus diseases. By far the most widely used in routine diagnostic work is complement fixation, because it is less dependent than are agglutination and precipitation upon the preparation of purified virus suspensions or extracts. In the rickettsial infections agglutination has now become a standard method and there are recent indications that it may eventually replace complement fixation for the serological diagnosis of some of the infections caused by the larger viruses, but the elementary bodies of most of the viruses which cause human diseases are too small to allow preparation of agglutinating suspensions without special apparatus beyond the reach of most diagnostic laboratories. Similarly, the extraction of suitable precipitating antigens usually presents difficulties that preclude their use for routine diagnosis, though there are a few notable exceptions, as for example in vaccinia and smallpox (Craigie & Tulloch, 1931).

Complement fixation, on the other hand, can now be used in the routine serological diagnosis of a very large number of virus infections of man. Many different techniques and modifications of the test had to be developed, however, to overcome the special difficulties presented by some of the diseases. Some viruses possess both group and specific complement-fixing antigens and it may be necessary in some cases to eliminate the group component and absorb out the corresponding antibody from a serum in order to achieve the necessary specific reaction (Bedson, Barwell, King & Bishop, 1949). Recent work suggests that the conglutination complement absorption test may on occasion be more sensitive than the orthodox haemolytic test (Stoker, Coombe & Bedson, 1950). A notable advance has been the introduction of micro-methods, which greatly facilitate large-scale serological surveys (Fulton & Dumbell, 1949; and see also Fulton, Plate I, fig. A, of this Bulletin).

In spite of these advances there still remain several important virus diseases in which complement fixation cannot be usefully applied, either because of the impracticability of obtaining antigenic material or because of the anticomplementary nature, or non-specific reactivity, of such material as is available. Brain tissue, the richest source of virus antigen in most of the encephalitides, is notoriously troublesome in complement-fixation tests, though the difficulties have been largely overcome in several cases by extraction techniques based upon the early researches of Craigie & Tulloch (1931). For example, Casals, Olitsky & Anslow (1951) have devised a method of obtaining a satisfactory poliomyelitis virus antigen from the homogenized brains of suckling mice infected with a rodent strain of the virus. Their success is partly attributable to the fact that brains of suckling mice contain much less lipoid material than brains of adults, and this introduces a new factor which is likely to be further exploited in other diseases for which serological diagnostic methods have hitherto been unavailable.

Two serological tests, namely, cold haemagglutination and *Streptococcus* M.G. agglutination, are useful in aiding the diagnosis of the atypical virus pneumonias (Turner, 1943; Peterson, Ham & Finland, 1943; Thomas, Mirick, Curnen, Ziegler & Horsfall, 1945). These are purely empirical tests and the elucidation of the scientific basis of the reactions involved must await the isolation and antigenic analysis of the causative viruses.

So far, no toxin separable from virus particles, analogous to the true bacterial exotoxins, has been demonstrated. Nevertheless, virus toxicity as distinct from infectivity is demonstrable with the viruses of the influenza and the psittacosis-lymphogranuloma groups (Rake & Jones, 1944; Henle & Henle, 1946a, 1946b). This toxic action is neutralizable by specific immune sera, and serum potency in this respect is closely correlated with its virus-neutralizing potency; probably the same antibodies are concerned in both reactions. Toxin neutralization has already been applied in the diagnosis of lymphogranuloma inguinale and is likely to find increasing diagnostic application in future, although at present its major interest and importance lies in connexion with the interactions between cell and virus involved in infection and host resistance.

2. Virus Haemagglutination

The phenomenon of virus haemagglutination was discovered independently by Hirst (1941) and by McClelland & Hare (1941). It consists in the fact that some viruses or virus derivatives unite with the red blood cells of certain animal species so that in suitable mixtures of red blood cell suspensions and virus the cells are linked together into aggregates, which then rapidly sediment. Haemagglutination has now been demonstrated with nearly a score of viruses and it is clear that the mechanism is not identical in all instances. The original observations were made with influenza virus[3], and the bulk of work since then has been done with viruses of the influenza group, but there can be little doubt that the complex mechanisms involved in the particular case of influenza have a wider significance. Discussion will therefore be confined to the reactions as they affect this group.

[2] See Downie & Macdonald, Plate IV, fig. J, of this Bulletin.

[3] See Isaacs, p. 208 of this Bulletin.

The reaction is specifically inhibited by homologous immune sera. Indeed, the specificity of the haemagglutination inhibition is extremely sharp, so that it can be employed for the differentiation of even closely related virus strains. So rapid, easy and convenient is it for the detection and assay of serum antibodies that for serological diagnosis and epidemiological surveys in influenza epidemics it has now virtually replaced all other methods. A fourfold rise in serum haemagglutinin inhibition titre during convalescence is generally accepted as establishing a retrospective diagnosis of influenza. The reaction is also invaluable for detecting the growth of virus in eggs inoculated with material from human cases, thus greatly facilitating the isolation and establishment of new strains. These can then be rapidly typed by the haemagglutination inhibition test. It thus may be possible in future to prepare prophylactic vaccines containing the current epidemic virus strains in time to meet the challenge of a major outbreak.

The implications of virus haemagglutination in respect of the mechanisms of infection and host resistance are possibly of greater importance than the practical applications. The attachment of virus to red cells is by means of specific linkage zones on the surfaces of virus and cell, conveniently referred to as virus and cell receptors respectively. It occurs with either living infective virus or with virus inactivated by heat or by various other means. With living virus the linkage is soon broken by the action of a virus enzyme that destroys the cell receptors so that the virus is freed to attach itself to fresh available cells. Those cells which have lost their receptors, however, are not susceptible to further interaction with virus. Although most work has been done with mixtures of virus and red blood cell suspensions, it has been shown that the fundamental reaction of union of virus and cell, followed by cell-receptor destruction and virus elution, pertains equally to cells of the respiratory tract of ferrets and mice, and to the cells lining the chorio-allantoic cavity of the embryonated egg; in other words, to cells which are susceptible to parasitization by virus during the course of infection. The work of Stone has been particularly valuable in this connexion. She showed that destruction of the receptors of cells lining the allantoic cavity of the chick embryo either abolishes or enormously reduces the susceptibility of the membrane to infection (Stone, 1948). It thus seems certain that receptor union with susceptible cells represents the first essential stage in virus infection. What part cell-receptor destruction by the virus enzyme plays in subsequent stages or in the spread of virus from cell to cell is unknown, but it has not been found possible to destroy this enzymic activity without also destroying the infectivity of the virus.

A number of biological fluids, including most normal sera, contain substances which inhibit virus haemagglutination. They are probably mucoproteins identical with or chemically related to the red-cell receptor substances and, like the latter, are rapidly destroyed by the virus enzyme. Because of this enzymic destruction the inhibitory action of the normal serum inhibitors is much more readily demonstrable if inactivated virus is used, and it is generally considered that on this account they can play little, if any, part in the resistance mechanisms of the host. Laboratory experiments have certainly failed to show them as possessing significant protective power against infection, but it has been suggested that they may represent merely the cast-off surface components of cells and that their protective efficacy whilst still an integral part of the susceptible cell may be considerable.

3. Antigenic Complexity of Viruses

The development of techniques for the preparation of relatively pure and concentrated virus suspensions, coupled with the increasingly successful application of in-vitro serological reactions to virus studies, has revealed that some of the larger viruses have a complexity of structure and of antigenic constitution comparable with that of bacteria. In this connexion the virus which has been most thoroughly investigated is that of vaccinia. Arising out of the important discovery of a soluble LS antigenic complex by Craigie & Wishart (1936), no less than four distinct serologically reactive surface antigens have now been separated and characterized. None of them, however, accounts for the development of virus-neutralizing and protective antibodies in the sera of infected human beings or animals, so that even in the case of vaccinia much further study is required to relate antigenic constitution and structure to the phenomena of infection and resistance.

There are clear indications that several other viruses likewise possess a multiplicity of antigens. It is not surprising, therefore, that antigenic components may be shared by viruses with similar biological activities. There are already indications that such relationships may eventually form one of the chief bases for virus classification. For example, virus strains within each of the major influenza virus types exhibit varying degrees of antigenic relationship, the viruses of psittacosis and lymphogranuloma inguinale give a considerable degree of serological cross-reaction, and similar cross-reactions are demonstrable between several members of the encephalitis group of viruses (Smith & Andrewes, 1938; Bedson, 1936; Barwell, 1948).

On the other hand, viruses without any demonstrable antigenic relationship may have almost identical biological activities. The two major influenza virus types A and B have for long been considered to give no cross-reactions in serological and immunological tests, though Smith (1952) claims that antigenic relationship based on a common deep antigen is demonstrable.

4. Antigenic Lability

In view of the antigenic complexity of many viruses, one might expect that antigenic variants, produced by natural selection of chance mutants, would arise from time to time. This process has in fact been directly observed on several occasions under laboratory conditions and the circumstantial evidence supporting the contention that it occurs frequently in nature is overwhelming. Antigenic lability, however, varies greatly with different viruses; some, like yellow fever virus, appear to be antigenically homogeneous in spite of variation of biological activity amongst strains, whilst others, like influenza, are exceptionally labile. Very soon after the first isolation of an influenza virus by Smith, Andrewes & Laidlaw in 1933, the possibility of strain differences being responsible for the short duration of immunity after infection was appreciated and the point was investigated as soon as sufficient strains were available for the purpose. By a number of cross-immunization experiments in both ferrets and mice and also by a long series of cross-serum neutralization tests, Smith & Andrewes (1938) were able to construct an antigenic spectrum covering the strains then available. Antigenic analysis was greatly facilitated by the introduction of the haemagglutination inhibition test and by complement-fixation reactions, so that representative strains from every major epidemic

occurring in different countries are now examined as a matter of routine investigation by the World Influenza Centre of the World Health Organization, at the National Institute for Medical Research, Mill Hill, London. The picture which emerges is one of constant evolution of new strains, so that the antigenic relationship between current and old strains tends to become more and more distant.

It is now well known that antigenic diversity amongst virus strains is not the main reason for the short duration of the immunity that follows recovery from influenza and some other kindred virus infections. It is, however, probably responsible for the occurrence of second attacks of those diseases that do engender solid immunity, such as poliomyelitis. It is also a factor of major importance in epidemiology. The building up of an effective herd immunity against an infective agent with the antigenic lability of influenza virus A is impossible, nor is it easy to see how such a disease can ever cease to be recurrently epidemic. Antigenic lability is also the chief factor militating against the production of effective prophylactic influenza vaccines. The most successful field vaccination trial on record was one in which the vaccine used happened to contain a strain antigenically identical with that which caused the subsequent epidemic, but the same vaccine was ineffective against other outbreaks (Salk, Menke & Francis, 1945; Rasmussen, Stokes & Smadel, 1947; Francis, Salk & Quilligan, 1947).

5. The Interference Phenomenon

The fact that one virus infection may afford some degree of protection against superimposed infection by another virus was recorded by Jenner (1804). The earliest scientific investigations of the phenomenon were concerned with plant pathology (McKinney, 1929); investigations pertaining to the animal viruses followed the observation of Hoskins (1935) that monkeys could be protected against a lethal dose of viscerotropic yellow fever virus by the simultaneous inoculation of a neurotropic virus variant. This observation was extended by Findlay & MacCallum (1937), who established two extremely important points: (i) that the protection could not be ascribed to early production of specific antibodies and, (ii) that the antigenically unrelated viruses of yellow fever and Rift Valley fever exhibited mutual interference.

Following these pioneer investigations, many workers studied the behaviour of pairs of viruses during experimental infections of susceptible animals or during their growth in chick embryos or in tissue cultures. A very large number of pairs, both antigenically related and antigenically unrelated, have been found by one or other of these techniques to exhibit interference. It is not always necessary that the interfering virus should be living, for there are several reports of interference by viruses that have been inactivated by heat or chemical treatment or ultra-violet light irradiation (Henle & Henle, 1943, 1944; Andrewes & Elford, 1947; Isaacs & Edney, 1950). This probably accounts for the fact that for virus cultivation small inocula often lead to better yields than do large inocula; presumably, if an inoculum contains a certain proportion of dead virus, multiplication of the living component may be inhibited by a process of auto-interference. Furthermore, interference is not an all-or-none phenomenon but may be either complete or partial, a fact which invites speculation as to its possible significance in the variability of clinical response during an epidemic.

The interference phenomenon thus represents a clear case of cellular immunity, for any participation by humoral factors is ruled out by its occurrence in embryonated eggs and tissue cultures which lack any antibody-forming mechanism. It is probable that different mechanisms are involved in different instances but that in all of them the cell change is due to some specific interaction between cell and interfering virus. Obvious possibilities are destruction or saturation of cell receptors, exhaustion of essential growth factors and deviation of enzyme systems. It is as yet far from clear how important virus interference is as a factor in host resistance against naturally acquired infections. Further elucidation of the phenomenon may indeed eventually provide new and effective methods of protection against some of the virus diseases for which orthodox immunization procedures have signally failed.

References

Andrewes, C. H. & Elford, W. J. (1947) *Brit. J. exp. Path.* **28,** 278

Barwell, C. F. (1948) *Nature, Lond.* **162,** 460

Bedson, S. P. (1936) *Brit. J. exp. Path.* **17,** 109

Bedson, S. P., Barwell, C. F., King, E. J. & Bishop, L. W. J. (1949) *J. clin. Path.* **2,** 241

Beveridge, W. I. B. & Burnet, F. M. (1946) *Spec. Rep. Ser. med. Res. Coun., Lond.* No. 256

Burnet, F. M. (1936) *Spec. Rep. Ser. med. Res. Coun., Lond.* No. 220

Casals, J., Olitsky, P. K. & Anslow, R. O. (1951) *J. exp. Med.* **94,** 123

Craigie, J. & Tulloch, W. J. (1931) *Spec. Rep. Ser. med. Res. Coun., Lond.* No. 156

Craigie, J. & Wishart, F. O. (1936) *J. exp. Med.* **64,** 803

Findlay, G. M. & MacCallum, F. O. (1937) *J. Path. Bact.* **44,** 405

Francis, T., jr, Salk, J. E. & Quilligan, J. J., jr (1947) *Amer. J. publ. Hlth,* **37,** 1013

Fulton, F. & Dumbell, K. R. (1949) *J. gen. Microbiol.* **3,** 97

Henle, G. & Henle, W. (1946a) *J. exp. Med.* **84,** 623

Henle, W. & Henle, G. (1946b) *J. exp. Med.* **84,** 639

Henle, W. & Henle, G. (1943) *Science,* **98,** 97

Henle, W. & Henle, G. (1944) *Amer. J. med. Sci.* **207,** 705

Hirst, G. K. (1941) *Science,* **94,** 22

Hoskins, M. (1935) *Amer. J. trop. Med.* **15,** 675

Isaacs, A. & Edney, M. (1950) *Aust. J. exp. Biol. med. Sci.* **28,** 219

Jenner, E. (1804) *Med. phys. J.* **12,** 97

Levaditi, C. (1926) *L'herpes et le zona.* Masson, Paris

Levaditi, C. (1946) *Précis de virologie médicale.* Masson, Paris

McClelland, L. & Hare, R. (1941) *Canad. publ. Hlth J.* **32,** 530

McKinney, H. H. (1929) *J. agric. Res.* **39,** 557

Peterson, O. L., Ham, T. M. & Finland, M. (1943) *Science,* **97,** 167

Rake, G. & Jones, H. P. (1944) *J. exp. Med.* **79,** 463

Rasmussen, A. F., Stokes, J. C. & Smadel, J. E. (1947) *Amer. J. Hyg.* **142,** 1942

Salk, J. E., Menke, W. J. & Francis, T. (1945) *Amer. J. Hyg.* **42,** 57

Smith, W. (1952) *Lancet,* **1,** 885

Smith, W. & Andrewes, C. H. (1938) *Brit. J. exp. Path.* **19,** 293

Smith, W., Andrewes, C. H. & Laidlaw, P. P. (1933) *Lancet,* **2,** 66

Stoker, M. G. P., Coombe, R. R. A. & Bedson, S. P. (1950) *Brit. J. exp. Path.* **31,** 217

Stone, J. D. (1948) *Aust. J. exp. Biol. med. Sci.* **26,** 49

Thomas, L., Mirick, G. S., Curnen, E. C., Ziegler, J. E., jr & Horsfall, F. L., jr (1945) *J. clin. Invest.* **24,** 227

Turner, J. C. (1943) *Nature, Lond.* **151,** 419

Woodruff, A. M. & Goodpasture, E. W. (1931) *Amer. J. Path.* **7,** 209

CHEMOTHERAPY OF VIRUS DISEASES

E. WESTON HURST D.Sc. M.D. F.R.C.P.

Biological Laboratories
Imperial Chemical Industries Limited
Manchester

1 The psittacosis-lymphogranuloma group of viruses
  a Experimental background
  b Therapy in man
2 The smaller viruses
3 Conclusion
References

For the present purpose we may divide infections caused by viruses into two categories. The first comprises maladies due to the large viruses of the psittacosis-lymphogranuloma group[1], against which it is relatively easy to find chemotherapeutic agents of greater or less potency. The second embraces all other virus diseases; here we can hardly do more than describe trends of chemotherapeutic research, and point to indications that the future may tell a more encouraging story. Within the limits imposed on this review, it is possible only to mention the salient facts relating to either group of infections; treatment is mainly factual and for a more theoretical approach to the subject the reader should consult the paper by Andrewes & King (1946).

1. The Psittacosis-Lymphogranuloma Group of Viruses

In addition to lymphogranuloma venereum and the ornithoses, this group includes trachoma, inclusion conjunctivitis and a number of diseases of animals, such as mouse pneumonitis and feline pneumonitis. Many regard primary atypical pneumonia as probably a virus disease.

a. Experimental Background

MacCallum & Findlay (1938), Bär (1938) and Levaditi (1938a, 1938b, 1938c, 1938d) demonstrated the therapeutic activity of the sulphonamides against lymphogranuloma in mice or guinea-pigs. Subsequent work of many authors established a general pattern of therapeutic behaviour common to the sulphonamides and to more recent and, in several instances, more potent drugs.

The different sulphonamides are not equally effective against lymphogranuloma in mice; sulphathiazole and sulphadiazine are among the more active (Findlay, 1940a; Jones, Rake & McKee, 1941; van den Ende & Lush, 1943). Most writers have failed to detect an action of the drugs in vitro (MacCallum & Findlay, 1938; Schlossberger & Bär, 1939; Rodaniche, 1943), or in tissue cultures in which neither cells nor virus are actively growing (Early & Morgan, 1946a). Although during therapy the titre of virus may be lower in the organs of treated than of control animals, virus is not completely eradicated and may persist in surviving mice for a year or more (Rodaniche, 1942, 1943; Jones, Rake & Stearns, 1945); doubtless, it is the immunity engendered during the period of therapy which prevents unrestricted growth of virus when dosing ceases (Hurst, Peters & Melvin, 1950). The action of the sulphonamides is inhibited, in part at any rate, by *p*-aminobenzoic acid (Findlay, 1940b; Mudrow & Bock, 1943; Rodaniche, 1943), and in tissue culture by pteroic and pteroylglutamic acids (Morgan, 1948) and by the citrovorum factor (Morgan, 1952a). A more detailed review of the treatment of experimental lymphogranuloma with sulphonamides is that of Jones *et al.* (1945).

The sulphonamides are also not equally effective against all strains of lymphogranuloma virus (Hurst *et al.* 1950) or against all viruses of the group. While the virus of mouse pneumonitis is susceptible to their action (Rake, Jones & Nigg, 1942; Rake & Hamre, 1944; Hamre & Rake, 1944), the last two authors failed to influence that of feline pneumonitis; Eaton & Hanford (1945) showed that the results of treatment depend to some extent on the host, and that the latter infection is more successfully treated in hamsters and rats than in chick embryos or mice. Rudd & Burnet (1941) failed to benefit psittacosis with several sulphonamides, but Meiklejohn, Wagner & Beveridge (1946), Wiseman, Meiklejohn, Lackman, Wagner & Beveridge (1946) and Early & Morgan (1946a, 1946b) found sulphadiazine effective against some American strains in both chick embryos and mice. Host differences were again apparent in the lower blood levels needed for successful therapy in the mouse as compared with the chick embryo.

Other synthetic chemical substances active in some degree or other against the largest viruses include acriflavine (Mauer, 1938), various nitroacridines (Eaton, van Allen & Wiener, 1947; Hurst, 1948), some substituted nitrobenzenes and nitrofurans (Eaton, Huang & Levenson, 1949), *p*-arsenobenzamide and its dithioglycollate (Eaton, Levenson & Hanks, 1949), and aromatic diamidines (Eaton, Levenson & Perry, 1952). We ourselves have observed slight or moderate activity in a number of quite different chemical types other than the above, and very marked activity in some quinoxaline-1:4-dioxides (Hurst, Landquist, Melvin, Peters, Senior, Silk & Stacey, 1953). The better quinoxaline oxides approximate to aureomycin in potency (see below); most, but not all, of them are more active against lymphogranuloma in the mouse than against lymphogranuloma in the chick embryo or against psittacosis in the mouse, once more illustrating the complex interrelationships between virus, host and chemotherapeutic agent. In tissue cultures a number of purine analogues and thiouracil, at concentrations not obviously toxic to the host cells, inhibit the growth of the 6-BC psittacosis virus (Morgan, 1952b, 1952c), and on exposure in vitro various enzyme-inhibitors reduce the activity of psittacosis virus (Burney & Golub, 1948).

Of the antibiotics, chloramphenicol, penicillin, aureomycin and terramycin all exhibit activity. Heilman & Herrell (1944a, 1944b) and Parker & Diefendorf (1944) demonstrated the susceptibility of ornithosis and psittacosis viruses to therapy with penicillin, and the drug was active, albeit at widely different doses, against each of 10 viruses of the psittacosis-lymphogranuloma group examined by Meiklejohn *et al.* (1946). Hamre & Rake (1947) found the agent of feline pneumonitis much less susceptible than that of lymphogranuloma, and Eaton, Dozois, van Allen, Parish & Schwalm (1948) found it less susceptible than the virus of murine

[1] See Bedson, p. 226 of this Bulletin.

pneumonitis. Against the American 6-BC strain of psittacosis virus, penicillin and sulphadiazine were roughly equal in activity, with the latter having the advantage when virus was introduced intracerebrally or intranasally (Meiklejohn *et al.* 1946; Wiseman, Meiklejohn, Lackman, Wagner & Beveridge, 1946; Early & Morgan, 1946a, 1946b). Against English strains of psittacosis, however, and against lymphogranuloma, massive doses of penicillin greatly surpass the sulphonamides in therapeutic potency (Bedson & May, 1945; Hurst *et al.* 1950). Smadel & Jackson (1948) equated chloramphenicol with penicillin and sulphadiazine, but in mice infected with either psittacosis or lymphogranuloma the results are disappointing if the animals are kept under observation for a sufficient time (Hurst *et al.* 1950). This drug is also unsatisfactory in the treatment of feline pneumonitis in the mouse (Kneeland & Price, 1950). Aureomycin, on the other hand, is highly effective (Wong & Cox, 1948; Wagner, 1949)—on a molecular basis five times more so than chloramphenicol (Wells & Finland, 1949). Moreover, unlike penicillin it is equally effective whatever the route of introduction of virus, and usually saves animals unless they are actually moribund when treatment begins; the activity of terramycin is very similar (Hurst *et al.* 1950).

The devotees of indiscriminate blunderbuss therapy may heed the warnings of Jawetz, Gunnison & Speck (1951), Jawetz, Gunnison, Speck & Coleman (1951) and Speck, Jawetz & Gunnison (1951) that at certain levels of dosage the antibacterial action of one antibiotic may be diminished by simultaneous exhibition of another; we have made analogous observations with lymphogranuloma virus (Hurst *et al.* 1950, 1953).

As stated above, the pattern of therapeutic behaviour with even the most potent drugs is similar to that with the sulphonamides. The quinoxaline oxides and the purest preparations of the antibiotics have little or no effect on the viruses in vitro. During therapy the titre of virus in the organs of treated animals drops, even to the point where virus is undetectable. Nevertheless, when dosing ceases virus reappears, and many animals remain carriers indefinitely. The exact proportion carrying virus depends partly on the level of dosage and partly on the drug employed (Hurst *et al.* 1950, 1953). There is evidence that, like the natural hosts, clinically recovered human beings also may carry psittacosis virus for long periods (up to 10 years—Meyer & Eddie, 1951); while lymphogranuloma virus has been isolated from inguinal nodes as long as 95 days after onset of the disease, though for not more than two or three weeks in patients treated with sulphathiazole (Heyman, Wall & Beeson, 1947).

The therapeutic action of penicillin is accompanied by morphological changes in the viruses, as seen in infected yolk sacs (Weiss, 1950; Gogolak & Weiss, 1950; Hurst *et al.* 1953). As well as invading considerably fewer cells, the viruses fail to divide into elementary bodies, and exist as large plaques up to 6 μ. in diameter. On the other hand, while reducing viral growth to a greater degree, aureomycin and the quinoxaline oxides do not inhibit division to the same extent as does penicillin.

b. Therapy in Man

Good agreement exists between the experimental data summarized above and the reports from the clinics.

Lymphogranuloma Venereum. This disease responds favourably to the sulphonamides, though chronic involvement of the vulva or rectal stricture may require months of treatment and the assistance of surgery. A convenient review is that by Koteen (1945). In early cases Willcox (1946) obtained excellent results with penicillin. Chloramphenicol (Woodward, 1949) has given good results in one early case. The most effective remedies, however, appear to be aureomycin and terramycin. Wright, Sanders, Logan, Prigot & Hill (1948), Prigot, Wright, Logan & de Luca (1949), Schamberg, Carrozzino & Boger (1951) and Fletcher, Sigel & Zintel (1951) bear witness to the virtues of the former, though Wammock, Greenblatt, Dienst, Chen & West (1950) consider the antibiotic more efficacious against the later manifestations of the disease. Wright, Whitaker, Wilkinson & Beinfield (1951) find terramycin similar in its action.

Psittacosis and the Ornithoses. These infections are more difficult to control by chemotherapy. Sulphonamides are of little or no value against most strains of virus (Meiklejohn, Beck & Eaton, 1944; Levinson, Gibbs & Beardwood, 1944; Olson & Treuting, 1944; Hamke, 1951), but penicillin has been moderately successful (Turgasen, 1944; Flippen, Gaydosh & Fittipoldi, 1945; Kirkwood, 1946; Rosebury, Ellingson, Meiklejohn & Schabel, 1947; Meyer & Eddie, 1947), except in the hands of Hamke (1951). Evidence available at the moment suggests that aureomycin exhibits still greater activity (Brainerd, Lennette, Meiklejohn, Bruyn & Clark, 1949; Green, 1950; Hamke, 1951).

Trachoma and Inclusion Conjunctivitis. Because of the relatively greater importance of secondary infections in the clinical course of trachoma, and the differences in severity of the disease in various parts of the world, the situation here is less easy to assess. By different observers, the sulphonamides have been considered ineffective in treatment, as useful adjuncts to therapy, or as curative agents in themselves; those with greatest experience of the disease tend to emphasize their value (Bellows, 1943; Thygeson, 1943; Poleff, 1943; Siniscal, 1952). Under treatment, the characteristic intracellular inclusions rapidly disappear (Thygeson, 1940), and in time infectivity also (Poleff, 1943; Sorsby, 1945), especially when both local and systemic treatment are combined with expression of the follicles on one or more occasions. Corneal lesions respond more readily than conjunctival, and papillary hypertrophy responds more readily than follicular hypertrophy (Thygeson, 1940). The drugs do not sterilize trachomatous tissue in vitro (Julianelle & Smith, 1942). Inclusion conjunctivitis also is beneficially affected by sulphonamides (Thygeson, 1943). Penicillin, on the other hand, has not proved of value in trachoma, and the early favourable report of Darius (1945) has not received confirmation (Lavery, 1946; Thygeson, 1948). Pijoan, Payne & Dineen (1950) claimed excellent results with chloramphenicol given systemically in trachoma. Several authors have been impressed by the effect of local or systemic aureomycin in trachoma or inclusion conjunctivitis (Braley & Sanders, 1949; Duke-Elder, Ainslie & Boase, 1950; Boase, 1950; Bellows, Richardson & Farmer, 1950), with others doubtful or dissenting as to its value against the trachomatous process itself (Shah, 1951; Sarkies, 1951; Ching, 1951). Terramycin similarly has been held to be of value (Mitsui, Tanaka, Toya, Iwashige & Yamashita, 1951; Cat, 1952). Applying different local treatments to the two eyes, Mitsui & Tanaka (1951) found terramycin to be superior to aureomycin in trachoma and both to be preferable to chloramphenicol. Aureomycin and terramycin were both active against inclusion blennorrhoea. Not all

the cases so treated were contaminated by bacteria, and as viral inclusions began to degenerate and disappear within 24 hours in acute cases the results suggest a direct action on the virus; the experimental evidence relating to the marked effect of these drugs on other viruses of the group renders this conclusion not unacceptable. Yet Siniscal (1952), from his experience with 3,500 patients, maintains that the antibiotics all give disappointing results and influence only the secondary invaders; according to him, the more highly soluble sulphonamides (sulphafurazole, sulphacetamide, sulphadiazine, etc.) constitute the first therapeutic choice, both because of their intrinsic activity and because of their blandness.

Primary Atypical Pneumonia. A number of facts support the view that the causative agent of primary atypical pneumonia may be a virus; it is generally agreed that sulphonamides and penicillin are not effective chemotherapeutic remedies for the condition. Although the varying severity of the cases makes assessment of the results of therapy far from simple, many clinical trials, some carried out with great care, have suggested a beneficial action of aureomycin in encouraging rapid defervescence and recovery (Brainerd *et al.* 1949; Finland, Collins & Wells, 1949; Kneeland, Rose & Gibson, 1949; Meiklejohn & Shragg, 1949; Schoenbach & Bryer, 1949; Anderson, 1950; Blodgett, Keating & Coffin, 1950; Cohen & Schwartz, 1950; Collins, Wells, Gocke & Finland, 1950; Schoenbach, Sweed, Tepper & Bryer, 1950; and others). Harvey, Mirick & Schaub (1949), however, failed to observe any improvement with aureomycin. In two patients Finland *et al.* (1949) claimed a therapeutic effect of streptomycin, Wood (1949) of chloramphenicol in one case and Melcher, Gibson, Rose & Kneeland (1950) of terramycin in seven cases. Comparing the results with three antibiotics, Graves & Ball (1951) found little to choose between aureomycin, chloramphenicol and terramycin on the score of therapeutic activity, but decided in favour of the last on the ground of lowest toxicity.

2. The Smaller Viruses

The antibiotics so successful in treating diseases caused by the largest viruses do not appear unequivocally to influence other virus infections. An exception to this statement concerns the effect of aureomycin and terramycin in curing clinically, *and eradicating virus from*, mice suffering from "grey lung disease", and cotton-rats from a similar condition (Andrewes & Niven, 1950a, 1950b). Andrewes (personal communication) also finds the virus susceptible to neoarsphenamine and tryparsamide.

In evaluating current experimental work, it is of the utmost importance to take into consideration the complex of host cell and virus that is studied. Many substances are capable of influencing the growth of bacteriophages, sometimes at concentrations which are not bacteriostatic (Spizizen, 1943a, 1943b; Jones, 1945; Fitzgerald & Babbit, 1946; Fitzgerald & Lee, 1946; Dickinson, 1948; Foster, 1948; Fowler & Cohen, 1948; Smith, 1949; Boyd & Bradley, 1951; Czekalowski, 1952; Fischer, 1952); among such substances acridines are prominent. As will be seen shortly, in tissue cultures also it is relatively easy to interfere with the multiplication of viruses, and recently attention has been paid to metabolic requirements for viral growth; there are those who believe that this work constitutes the "rational approach" to the chemotherapy of virus diseases. It is less easy to influence the growth of most viruses in the chick embryo, and protective effects of any magnitude in the mouse or hatched chick are rare. My former colleague, Mr J. Francis, injected many hundreds of compounds into the yolk sac of chick embryos and found that a score or so were significantly active in restricting the growth of fowl-pox virus on the chorioallantoic membrane. None of these produced the slightest effect on fowl-pox in the hatched bird. A number of nitroacridines and some chloroacridines reduce growth of several viruses in tissue culture or in the chick embryo (Green, Rasmussen & Smadel, 1946; Briody & Stannard, 1951; Eaton, Cheever & Levenson, 1951; Rasmussen & Stokes, 1951), but most compounds of this type are not appreciably active in the mouse. On the other hand, mepacrine, which in the mouse shows marked activity against equine encephalomyelitis, and less activity against several other infections, is inactive against equine encephalomyelitis in the chick embryo (Hurst, Melvin & Peters, 1952; Hurst, Peters & Melvin, 1952). Therefore, while studies on bacteriophages or tissue cultures may help in solving the basic problems connected with viral growth and reproduction, it cannot be affirmed with conviction that they will lead to practical chemotherapy sooner than the partially empirical screening of compounds in the mouse carried out in some commercial laboratories, and on a strictly limited scale elsewhere (Coggeshall & Maier, 1942; Andrewes, King & van den Ende, 1943; Krueger *et al.* 1943; Kramer, Geer & Szobel, 1944; Cutting, Dreisbach, Halpern, Irwin, Jenkins, Proescher & Tripi, 1947).

Many of the substances examined in tissue cultures have been enzyme-inhibitors, metabolites, or the analogues of metabolites. Thus Thompson (1947) found cyanide, azide, atabrine, proflavine, iodoacetic acid and other enzyme-inhibitors, certain substituted amino-acids and benzimidazole inhibitory to the vaccinia virus. Subsequent work has followed the same general lines and the data in Table I are representative of those which have accumulated. In many instances inhibition of growth is reversed by related metabolites, or by other substances.

Effects in the mouse have usually been slight, of dubious significance, or wanting, even though near-toxic doses of drug have been administered from approximately the moment of infection. However, the following observations seem worthy of mention.

While 2:6-diaminopurine and benzimidazole do not protect mice against intracerebrally inoculated vaccinia (Thompson, Price, Minton, Elion & Hitchings, 1950), nor the former against West Nile, Ilhéus or louping-ill viruses (Moore & Friend, 1951), the latter writers demonstrated a protective effect of the aminopurine against invasion of the brain in Russian spring-summer encephalitis. Certain 5-phenoxythiouracils, only moderately active in tissue culture, protected against intracerebral vaccinia but not against influenza, herpes simplex, St Louis, MM or Rift Valley fever viruses (Thompson, Price, Minton, Falco & Hitchings, 1951). Thiosemicarbazones have been reported active against vaccinia virus given intracerebrally or intranasally (Thompson, Price & Minton, 1951; Hamre, Brownlee & Donovick, 1951). Malononitrile preserves the life of many mice infected with low concentrations of the Lansing poliomyelitis virus (Szanto & Felsenfeld, 1949). Some pyrimidines may exert a slight action on infections with Columbia-SK virus (McKinstry & Reading, 1944), as do certain naphthoquinonimines (Jungeblut, 1951; Schnitzer, Buck & Steiger,

TABLE I. SUBSTANCES INHIBITING THE GROWTH OF THE SMALLER VIRUSES IN TISSUE CULTURE

Substance	Virus	Author
Various substituted amino-acids	Vaccinia	Thompson (1947)
β-2-Thienylalanine	Vaccinia	Thompson & Wilkin (1948)
Arginine, lysine, ornithine	Influenza, mumps	Eaton, Magasanik, Perry & Karibian (1951)
DL-Ethionine	Poliomyelitis	Brown & Ackermann (1951)
DL-Methoxinine, DL-ethionine	Influenza	Ackermann (1951a)
Various analogues of amino-acids	Poliomyelitis	Brown (1952)
L-Lysine and other amino-acids	Mouse encephalitis (GD VII)	Pearson, Lagerborg & Winzler (1952)
o-, *m*-, or *p*-Fluorophenylalanine	GD VII	Pearson *et al.* (1952)
Various aminosulphonic acids	Influenza	Ackermann (1952)
Benzimidazole	Vaccinia	Thompson (1947); Thompson *et al.* (1950)
Various halogenated purines	Vaccinia	Thompson *et al.* (1950)
2:6-Diaminopurine	Vaccinia	Thompson *et al.* (1950)
2:6-Diaminopurine	Russian spring-summer encephalitis	Friend (1951)
2:6-Diaminopurine	Poliomyelitis	Brown (1952)
Benzimidazole	Poliomyelitis	Brown (1952)
2:5-Dimethylbenzimidazole	Influenza	Tamm, Folkers & Horsfall (1952)
Proflavine	Vaccinia	Thompson (1947)
Proflavine	Poliomyelitis	Brown (1952)
Glycerin monoacetin	Poliomyelitis	Brown (1952)
Oxythiamine, deoxypyridoxine	Influenza, mumps	Cushing & Morgan (1952)
DL-Homobiotin	Poliomyelitis	Brown (1952)
5-Chloroacetamidopyrimidine, etc.	Vaccinia	Thompson, Wilkin, Hitchings & Russell (1949)
Diamidines	Influenza, mumps	Eaton, Perry, Levenson & Gocke (1952)
Cobaltous ion	Influenza	Schmidt & Rasmussen (1952)
Benzaldehyde thiosemicarbazone	Vaccinia	Thompson, Price & Minton (1951)

1951). Two organic mercurial compounds partially suppress the excretion of Theiler's virus (LoGrippo, Earle, Brodie, Graef, Bowman & Ward, 1952; LoGrippo, Earle, Graef & Ward, 1952). A dose of fluoroacetate interferes with the oxidation of citric acid and suppresses or retards for a time the growth of influenza and mouse-adapted poliomyelitis viruses (Ackermann, 1951b; Ainslie, 1952), but only at a level lethal for some animals (Mogabgab & Horsfall, 1952).

Claims to have benefited influenza in the mouse include those of Seeler, Graessle & Ott (1946) with quinine, Rubin & Giarman (1947) with certain lactones and their analogues, and Fleisher (1949) with Janus green B. Our own experience has shown that in murine influenza the significance of any results unaccompanied by daily titrations of virus is open to question. We ourselves have noted an increased mean period of survival and on occasion slight but reproducible decrease in mortality, with a number of wholly unrelated compounds administered in doses which permit normal or nearly normal growth of the treated animals; none of these treatments has restricted viral growth in the slightest degree, and we can only conclude that the effect has been upon the highly peculiar tissue reaction that follows growth of virus in the mouse and leads to death from asphyxia. An interesting bacterial product with precisely this property has recently been described by Groupé, Pugh & Levine (1952). On the other hand, the bacterial polysaccharides studied by Horsfall & McCarty (1947) and by Ginsberg, Goebel & Horsfall (1948) in virus pneumonia of mice and in mumps in the chick embryo apparently did restrict viral growth, even when introduced after virus had penetrated the susceptible cells.

Gershoff, Rasmussen, Elvehjem & Clark (1952) have drawn attention to the protective effect of feeding 6-methyltryptophan and excess methionine to mice infected with mouse-adapted poliomyelitis, an observation recalling that of Sprunt (1942) in reducing with methionine the severity of vaccinia in rabbits. Space will not permit further mention of the effect of nutritional factors on susceptibility to viruses; the data were reviewed by Clark (1950) and have since been supplemented by Davies, Pond, Smith, Rasmussen, Elvehjem & Clark (1952) and Pond, Davies, Smith, Elvehjem, Rasmussen & Clark (1952).

Sulkin, Zarafonetis & Goth (1946) recorded the protective effect of long periods of ether anaesthesia in mice infected with Eastern and Western equine and St Louis encephalomyelitis viruses, but not with rabies or mouse-adapted poliomyelitis. The first three belong to the group of viruses readily inactivated in vitro by ether, the last two to the group of ether-resistant viruses (Sulkin & Zarafonetis, 1947; Andrewes & Horstmann, 1949). The concentration of ether likely to be achieved, even locally, in the tissues is far below that needed for inactivation in vitro, a fact for which a number of possible explanations suggest themselves.

Starting from the observations of Wood & Rusoff (1945) that vital staining with trypan red diminishes susceptibility of mice to MM virus injected intraperitoneally, Hurst, Melvin & Peters (1952) and Hurst, Peters & Melvin (1952) discovered many substances modifying infections with equine encephalomyelitis virus; of these substances, mepacrine showed marked therapeutic properties. The drug was rather less active in Rift Valley fever, louping ill and herpes simplex, and detectably active in lymphocytic choriomeningitis and St Louis encephalitis, but not in 16 other virus infections. From this and other work on bacteriophage and in tissue culture (cited above), Hurst and his colleagues concluded that many acridines possess antiviral properties, but that in addition mepacrine

possesses pharmacological properties allowing its antiviral action to be manifested in the mouse; this behaviour may be related to the high concentrations it attains in some of the organs.

3. Conclusion

Thus we see that the largest viruses of the psittacosis-lymphogranuloma group are readily susceptible to chemotherapy, and that the facts known about them can already be woven into a consecutive narrative. On the other hand, consideration of the chemotherapy of diseases caused by the smaller viruses is no more than a description of trends of research and a string of unconnected observations of various degrees of significance. Nevertheless, indications are not wanting that these diseases may ultimately be susceptible to modification by chemical agency.

REFERENCES

Ackermann, W. W. (1951a) *J. exp. Med.* **93,** 337
Ackermann, W. W. (1951b) *J. exp. Med.* **93,** 635
Ackermann, W. W. (1952) *Proc. Soc. exp. Biol., N.Y.* **80,** 362
Ainslie, J. D. (1952) *J. exp. Med.* **95,** 9
Anderson, C. E. (1950) *Amer. J. Dis. Child.* **80,** 533
Andrewes, C. H. & Horstmann, D. M. (1949) *J. gen. Microbiol.* **3,** 290
Andrewes, C. H. & King, H. (1946) *Brit. med. Bull.* **4,** 931
Andrewes, C. H., King, H. & Ende, M. van den (1943) *J. Path. Bact.* **55,** 173
Andrewes, C. H. & Niven, J. S. F. (1950a) *Brit. J. exp. Path.* **31,** 767
Andrewes, C. H. & Niven, J. S. F. (1950b) *Brit. J. exp. Path.* **31,** 773
Bär, F. (1938) *Klin. Wschr.* **17,** 588
Bedson, S. P. & May, H. P. (1945) *Lancet,* **2,** 394
Bellows, J. G. (1943) *Arch. Ophthal., N.Y.* **29,** 888
Bellows, J. G., Richardson, V. M. & Farmer, C. J. (1950) *Amer. J. Ophthal.* **33,** 273
Blodgett, W. A., Keating, J. H. & Coffin, G. J. (1950) *J. Amer. med. Ass.* **143,** 878
Boase, A. J. (1950) *Brit. J. Ophthal.* **34,** 627
Boyd, J. S. K. & Bradley, P. L. (1951) *Brit. J. exp. Path.* **32,** 397
Brainerd, H., Lennette, E. H., Meiklejohn, G., Bruyn, H. B. & Clark, W. H. (1949) *J. clin. Invest.* **28,** 992
Braley, A. E. & Sanders, M. (1949) *Amer. J. Ophthal.* **32,** 119
Briody, B. A. & Stannard, C. (1951) *J. Immunol.* **67,** 423
Brown, G. C. (1952) *J. Immunol.* **69,** 441
Brown, G. C. & Ackermann, W. W. (1951) *Proc. Soc. exp. Biol., N.Y.* **77,** 367
Burney, T. E. & Golub, O. J. (1948) *J. Immunol.* **60,** 213
Cat, N. D. (1952) *Amer. J. trop. Med. Hyg.* **1,** 799
Ching, R. (1951) *Arch. Ophthal., N.Y.* **45,** 657
Clark, P. F. (1950) *Ann. Rev. Microbiol.* **4,** 343
Coggeshall, L. T. & Maier, J. (1942) *J. Pharmacol.* **76,** 161
Cohen, P. & Schwartz, R. (1950) *J. Pediat.* **37,** 23
Collins, H. S., Wells, E. B., Gocke, T. M. & Finland, M. (1950) *Amer. J. Med.* **8,** 4
Cushing, R. T. & Morgan, H. R. (1952) *Proc. Soc. exp. Biol., N.Y.* **79,** 497
Cutting, W. C., Dreisbach, R. H., Halpern, R. M., Irwin, E. A., Jenkins, D. W., Proescher, F. & Tripi, H. B. (1947) *J. Immunol.* **57,** 379
Czekalowski, J. W. (1952) *Brit. J. exp. Path.* **33,** 57
Darius, D. J. (1945) *Amer. J. Ophthal.* **28,** 1007
Davies, W. L., Pond, W. L., Smith, S. C., Rasmussen, A. F., Elvehjem, C. A. & Clark, P. F. (1952) *J. Bact.* **64,** 571
Dickinson, L. (1948) *J. gen. Microbiol.* **2,** 154
Duke-Elder, S., Ainslie, D. & Boase, A. J. (1950) *Brit. J. Ophthal.* **34,** 30
Early, R. L. & Morgan, H. R. (1946a) *J. Immunol.* **53,** 151
Early, R. L. & Morgan, H. R. (1946b) *J. Immunol.* **53,** 251
Eaton, M. D., Allen, A. van & Wiener, A. (1947) *Proc. Soc. exp. Biol., N.Y.* **66,** 141
Eaton, M. D., Cheever, F. S. & Levenson, C. G. (1951) *J. Immunol.* **66,** 463
Eaton, M. D., Dozois, T. F., Allen, A. van, Parish, V. L. & Schwalm, S. (1948) *J. Immunol.* **58,** 251
Eaton, M. D. & Hanford, V. L. (1945) *Proc. Soc. exp. Biol., N.Y.* **59,** 63
Eaton, M. D., Huang, C. & Levenson, C. G. (1949) *Proc. Soc. exp. Biol., N.Y.* **71,** 501
Eaton, M. D., Levenson, C. G. & Hanks, C. (1949) *Proc. Soc. exp. Biol., N.Y.* **71,** 508
Eaton, M. D., Levenson, C. G. & Perry, M. (1952) *J. Immunol.* **68,** 1
Eaton, M. D., Magasanik, B., Perry, M. E. & Karibian, D. (1951) *Proc. Soc. exp. Biol., N.Y.* **77,** 505
Eaton, M. D., Perry, M. E., Levenson, C. G. & Gocke, I. M. (1952) *J. Immunol.* **68,** 321
Ende, M. van den & Lush, D. (1943) *J. Path. Bact.* **55,** 81
Findlay, G. M. (1940a) *Lancet,* **2,** 682
Findlay, G. M. (1940b) *Brit. J. exp. Path.* **21,** 356
Finland, M., Collins, H. S. & Wells, E. B. (1949) *New Engl. J. Med.* **240,** 241
Fischer, G. (1952) *Acta path. microbiol. scand.* **31,** 433
Fitzgerald, R. J. & Babbit, D. (1946) *J. Immunol.* **52,** 121
Fitzgerald, R. J. & Lee, M. E. (1946) *J. Immunol.* **52,** 127
Fleisher, M. S. (1949) *J. Immunol.* **62,** 245
Fletcher, A., Sigel, M. M. & Zintel, H. A. (1951) *Arch. Surg., Chicago,* **62,** 239
Flippen, H. F., Gaydosh, M. J. & Fittipoldi, W. V. (1945) *J. Amer. med. Ass.* **128,** 280
Foster, R. A. C. (1948) *J. Bact.* **56,** 795
Fowler, C. B. & Cohen, S. S. (1948) *J. exp. Med.* **87,** 259
Friend, C. (1951) *Proc. Soc. exp. Biol., N.Y.* **78,** 150
Gershoff, S. N., Rasmussen, A. F., Elvehjem, C. A. & Clark, P. F. (1952) *Proc. Soc. exp. Biol., N.Y.* **81,** 484
Ginsberg, H. S., Goebel, W. F. & Horsfall, F. L. (1948) *J. exp. Med.* **87,** 385
Gogolak, F. M. & Weiss, E. (1950) *J. infect. Dis.* **87,** 264
Graves, F. B. & Ball, W. O. (1951) *J. Pediat.* **39,** 155
Green, R. H., Rasmussen, A. F. & Smadel, J. E. (1946) *Publ. Hlth Rep., Wash.* **61,** 1401
Green, T. W. (1950) *J. Amer. med. Ass.* **144,** 237
Groupé, V., Pugh, L. H. & Levine, A. S. (1952) *Proc. Soc. exp. Biol., N.Y.* **80,** 710
Hamke, H. (1951) *Dtsch. med. Wschr.* **76,** 246
Hamre, D., Brownlee, K. A. & Donovick, R. (1951) *J. Immunol.* **67,** 305
Hamre, D. & Rake, G. (1944) *J. infect. Dis.* **74,** 206
Hamre, D. & Rake, G. (1947) *J. infect. Dis.* **81,** 175
Harvey, J. C., Mirick, G. S. & Schaub, I. G. (1949) *J. clin. Invest.* **28,** 987
Heilman, F. R. & Herrell, W. E. (1944a) *Proc. Mayo Clin.* **19,** 57
Heilman, F. R. & Herrell, W. E. (1944b) *Proc. Mayo Clin.* **19,** 204
Heyman, A., Wall, M. J. & Beeson, P. B. (1947) *Amer. J. Syph.* **31,** 81
Horsfall, F. L. & McCarty, M. (1947) *J. exp. Med.* **85,** 623
Hurst, E. W. (1948) *Brit. J. Pharmacol.* **3,** 181
Hurst, E. W., Landquist, J. K., Melvin, P., Peters, J. M., Senior, N., Silk, J. A. & Stacey, G. J. (1953) *Brit. J. Pharmacol.* **8,** 297
Hurst, E. W., Melvin, P. & Peters, J. M. (1952) *Brit. J. Pharmacol.* **7,** 455
Hurst, E. W., Peters, J. M. & Melvin, P. (1950) *Brit. J. Pharmacol.* **5,** 611
Hurst, E. W., Peters, J. M. & Melvin, P. (1952) *Brit. J. Pharmacol.* **7,** 473
Jawetz, E., Gunnison, J. B. & Speck, R. S. (1951) *Amer. J. med. Sci.* **222,** 404
Jawetz, E., Gunnison, J. B., Speck, R. S. & Coleman, V. R. (1951) *Arch. intern. Med.* **87,** 349
Jones, D. (1945) *J. Bact.* **50,** 341
Jones, H. P., Rake, G. & McKee, C. N. (1941) *Proc. Soc. exp. Biol., N.Y.* **48,** 318
Jones, H. P., Rake, G. & Stearns, B. (1945) *J. infect. Dis.* **76,** 55
Julianelle, L. A. & Smith, J. E. (1942) *Amer. J. Ophthal.* **25,** 317

Jungeblut, C. W. (1951) *Proc. Soc. exp. Biol., N.Y.* **77,** 176
Kirkwood, T. (1946) *Illinois med. J.* **90,** 193
Kneeland, Y. & Price, K. M. (1950) *J. Immunol.* **65,** 653
Kneeland, Y., Rose, H. M. & Gibson, C. D. (1949) *Amer. J. Med.* **6,** 41
Koteen, H. (1945) *Medicine, Baltimore,* **24,** 1
Kramer, S. D., Geer, H. A. & Szobel, D. A. (1944) *J. Immunol.* **49,** 273
Krueger, A. P. *et al.* (1943) *Science,* **98,** 348
Lavery, F. S. (1946) *Brit. J. Ophthal.* **30,** 591
Levaditi, C. (1938a) *C. R. Soc. Biol., Paris,* **127,** 958
Levaditi, C. (1938b) *C. R. Soc. Biol., Paris,* **128,** 138
Levaditi, C. (1938c) *C. R. Soc. Biol., Paris,* **128,** 875
Levaditi, C. (1938d) *C. R. Soc. Biol., Paris,* **129,** 490
Levinson, D. C., Gibbs, J. & Beardwood, J. T. (1944) *J. Amer. med. Ass.* **126,** 1079
LoGrippo, G. A., Earle, D. P., Brodie, B. B., Graef, I., Bowman, R. & Ward, R. (1952) *Amer. J. Hyg.* **55,** 55
LoGrippo, G. A., Earle, D. P., Graef, I. & Ward, R. (1952) *Amer. J. Hyg.* **55,** 70
MacCallum, F. O. & Findlay, G. M. (1938) *Lancet,* **2,** 136
McKinstry, D. W. & Reading, E. H. (1944) *J. Franklin Inst.* **237,** 432
Mauer, G. (1938) *Zbl. Bakt.* (1 Abt., Orig.) **142,** 279
Meiklejohn, G., Beck, M. D. & Eaton, M. D. (1944) *J. clin. Invest.* **23,** 167
Meiklejohn, G. & Shragg, R. I. (1949) *J. Amer. med. Ass.* **140,** 391
Meiklejohn, G., Wagner, J. C. & Beveridge, G. W. (1946) *J. Immunol.* **54,** 1
Melcher, G. W., Gibson, C. D., Rose, H. M. & Kneeland, Y. (1950) *J. Amer. med. Ass.* **143,** 1303
Meyer, K. F. & Eddie, B. (1947) *J. Amer. med. Ass.* **133,** 822
Meyer, K. F. & Eddie, B. (1951) *J. infect. Dis.* **88,** 109
Mitsui, Y. & Tanaka, C. (1951) *Antibiot. Chemother.* **1,** 146
Mitsui, Y., Tanaka, C., Toya, H., Iwashige, Y. & Yamashita, K. (1951) *Arch. Ophthal., N.Y.* **46,** 235
Mogabgab, W. J. & Horsfall, F. L. (1952) *J. exp. Med.* **96,** 531
Moore, A. E. & Friend, C. (1951) *Proc. Soc. exp. Biol., N.Y.* **78,** 153
Morgan, H. R. (1948) *J. exp. Med.* **88,** 285
Morgan, H. R. (1952a) *J. exp. Med.* **95,** 269
Morgan, H. R. (1952b) *J. exp. Med.* **95,** 277
Morgan, H. R. (1952c) *Proc. Soc. exp. Biol., N.Y.* **81,** 448
Mudrow, L. & Bock, M. (1943) *Z. ImmunForsch.* **104,** 463
Olson, B. J. & Treuting, W. L. (1944) *Publ. Hlth Rep., Wash.* **59,** 1299
Parker, R. F. & Diefendorf, H. W. (1944) *Proc. Soc. exp. Biol., N.Y.* **57,** 351
Pearson, H. E., Lagerborg, D. L. & Winzler, R. J. (1952) *Proc. Soc. exp. Biol., N.Y.* **79,** 409
Pijoan, M. J., Payne, E. H. & Dineen, J. (1950) *Amer. J. trop. Med.* **30,** 677
Poleff, L. (1943) *Ophthalmologica, Basel,* **105,** 156
Pond, W. L., Davies, W. L., Smith, S. C., Elvehjem, C. A., Rasmussen, A. F. & Clark, P. F. (1952) *J. Bact.* **64,** 583
Prigot, A., Wright, L. T., Logan, M. A. & Luca, F. R. de (1949) *N.Y. St. J. Med.* **49,** 1911
Rake, G. & Hamre, D. M. (1944) *Proc. Soc. exp. Biol., N.Y.* **55,** 90
Rake, G., Jones, H. & Nigg, C. (1942) *Proc. Soc. exp. Biol., N.Y.* **49,** 449
Rasmussen, A. F. & Stokes, J. C. (1951) *J. Immunol.* **66,** 237
Rodaniche, E. C. (1942) *J. infect. Dis.* **70,** 58
Rodaniche, E. C. (1943) *J. infect. Dis.* **73,** 173
Rosebury, T., Ellingson, H. V., Meiklejohn, G. & Schabel, F. (1947) *J. infect. Dis.* **80,** 64
Rubin, B. A. & Giarman, N. J. (1947) *Yale J. Biol. Med.* **19,** 1017
Rudd, G. V. & Burnet, F. M. (1941) *Aust. J. exp. Biol. med. Sci.* **19,** 33
Sarkies, J. W. R. (1951) *Brit. J. Ophthal.* **35,** 549
Schamberg, I. L., Carrozzino, O. M. & Boger, W. P. (1951) *Amer. J. Syph.* **35,** 370
Schlossberger, H. & Bär, F. (1939) *Zbl. Bakt.* (1 Abt., Orig.) **144,** 228
Schmidt, J. R. A. & Rasmussen, A. F. (1952) *Proc. Soc. exp. Biol., N.Y.* **81,** 244
Schnitzer, R. J., Buck, M. & Steiger, N. (1951) *Proc. Soc. exp. Biol., N.Y.* **77,** 182
Schoenbach, E. B. & Bryer, M. S. (1949) *J. Amer. med. Ass.* **139,** 275
Schoenbach, E. B., Sweed, A., Tepper, B. & Bryer, M. S. (1950) *New Engl. J. Med.* **243,** 799
Seeler, A. O., Graessle, O. & Ott, W. H. (1946) *J. infect. Dis.* **79,** 156
Shah, M. A. (1951) *Brit. J. Ophthal.* **35,** 50
Siniscal, A. A. (1952) *J. Amer. med. Ass.* **148,** 637
Smadel, J. E. & Jackson, E. B. (1948) *Proc. Soc. exp. Biol., N.Y.* **67,** 478
Smith, A. G. (1949) *J. Immunol.* **61,** 57
Sorsby, A. (1945) *Brit. J. Ophthal.* **29,** 98
Speck, R. S., Jawetz, E. & Gunnison, J. B. (1951) *Arch. intern. Med.* **88,** 168
Spizizen, J. (1943a) *J. infect. Dis.* **73,** 212
Spizizen, J. (1943b) *J. infect. Dis.* **73,** 222
Sprunt, D. H. (1942) *Proc. Soc. exp. Biol., N.Y.* **51,** 226
Sulkin, S. E. & Zarafonetis, C. (1947) *J. exp. Med.* **85,** 559
Sulkin, S. E., Zarafonetis, C. & Goth, A. (1946) *J. exp. Med.* **84,** 277
Szanto, P. B. & Felsenfeld, P. (1949) *Proc. Soc. exp. Biol., N.Y.* **72,** 15
Tamm, I., Folkers, K. & Horsfall, F. L. (1952) *Yale J. Biol. Med.* **24,** 559
Thompson, R. L. (1947) *J. Immunol.* **55,** 345
Thompson, R. L., Price, M. L. & Minton, S. A., jr (1951) *Proc. Soc. exp. Biol., N.Y.* **78,** 11
Thompson, R. L., Price, M. L., Minton, S. A., jr, Elion, G. B. & Hitchings, G. H. (1950) *J. Immunol.* **65,** 529
Thompson, R. L., Price, M., Minton, S. A., jr, Falco, E. A. & Hitchings, G. H. (1951) *J. Immunol.* **67,** 483
Thompson, R. L. & Wilkin, M. L. (1948) *Proc. Soc. exp. Biol., N.Y.* **68,** 434
Thompson, R. L., Wilkin, M. L., Hitchings, G. H. & Russell, P. B. (1949) *Proc. Soc. exp. Biol., N.Y.* **72,** 169
Thygeson, P. (1940) *Amer. J. Ophthal.* **23,** 679
Thygeson, P. (1943) *Amer. J. Ophthal.* **29,** 1000
Thygeson, P. (1948) In: Rivers, T. M., ed. *Viral and rickettsial infections of man.* Lippincott, London
Turgasen, F. E. (1944) *J. Amer. med. Ass.* **126,** 1150
Wagner, J. C. (1949) *J. clin. Invest.* **28,** 1049
Wammock, V. S., Greenblatt, R. B., Dienst, R. B., Chen, C. & West, R. M. (1950) *J. invest. Derm.* **14,** 427
Weiss, E. (1950) *J. infect. Dis.* **87,** 249
Wells, E. B. & Finland, M. (1949) *Proc. Soc. exp. Biol., N.Y.* **72,** 365
Willcox, R. R. (1946) *Post Grad. med. J.* **22,** 97
Wiseman, R. W., Meiklejohn, G., Lackman, D. B., Wagner, J. C. & Beveridge, G. W. (1946) *J. Immunol.* **54,** 9
Wong, S. C. & Cox, H. R. (1948) *Ann. N.Y. Acad. Sci.* **51,** 290
Wood, E. J. (1949) *Lancet,* **2,** 55
Wood, H. G. & Rusoff, I. I. (1945) *J. exp. Med.* **82,** 297
Woodward, T. E. (1949) *Ann. intern. Med.* **31,** 53
Wright, L. T., Sanders, M., Logan, M. A., Prigot, A. & Hill, L. M. (1948) *J. Amer. med. Ass.* **138,** 408
Wright, L. T., Whitaker, J. C., Wilkinson, R. S. & Beinfield, M. S. (1951) *Antibiot. Chemother.* **1,** 193

TECHNIQUES FOR THE STUDY OF VIRUSES

FORREST FULTON D.M.

London School of Hygiene and Tropical Medicine
University of London

1 Properties of the virus elementary body
  a Size
  b Shape
  c Structure
  d Chemical composition
2 Mechanism of adsorption to the cell
3 Enumeration of elementary bodies
4 Purification
5 Storage
6 Multiplication of virus inside the cell
7 Methods of cultivation
8 Conclusions
References

In the hierarchy of those who work with viruses it is customary to allot a very humble place to the craftsmen who concern themselves with the minutiae of technique: for this reason it is necessary to mould the introductory section of an essay on methods into the form of an apologia.

The study of viruses is an experimental science because it is impracticable to observe them in unsophisticated ways. Direct study is made possible only by an elaborate instrumental extension of our senses, and if viruses are studied indirectly by the observation of effects, their properties are discovered only by a complicated series of inferences. Thus the technique of the experiment is the rock on which the whole edifice of the science is built.

New techniques are devised in an effort to answer problems which cannot conveniently, or with the necessary precision, be answered by the use of existing techniques. Once a new method has been evolved it may provide the stimulus for the study of other problems which were not, until then, precisely stated. Indeed, an experimental science is most nobly served by the continual interaction of new hypotheses with new techniques.

1. Properties of the Virus Elementary Body

a. Size

By 1925 it had been established that viruses which infect the cells of animals could pass through filter candles made of porcelain or kieselguhr which would retain small bacteria, the diameter of which was known by direct measurement under the microscope. It was also known that, whereas some viruses would traverse such filters very readily, others did so only with the greatest difficulty. These two facts suggested that viruses were less than 500 mμ. in diameter and probably differed greatly in size. More than this could not be inferred from the results of the filtration methods then available. At the time, there were two types of gel membrane which gave promise of providing a graded series of filters with different pore sizes. Bechhold (1907) had developed a series of membranes, prepared from acetic acid solutions of collodion: by varying the initial concentration of the solution it was possible to obtain membranes the average pore diameter of which was within the range of 50–1,000 mμ. These membranes were weak and required a filter paper support, but their serious disadvantage was the wide variation in pore diameter found in a single membrane; the ratio of the largest pore diameter to the average pore diameter was of the order of 10.

Another type of gel membrane was prepared from ethanol-ether solutions of collodion (Zsigmondy & Bachmann, 1918). They were much stronger and needed no support but it was difficult to obtain membranes with an average pore diameter greater than 100 mμ.

In 1927, W. J. Elford began a study of gel membrane filters. He investigated in particular the more promising ethanol-ether type and he observed that the aggregation of the collodion at the moment of gelation was a decisive factor in determining the size of the pores of the membrane. By adding to the ethanol-ether solution suitable amounts of acetone and amyl alcohol he was able to prepare a solution (the parent solution) which, when poured into a glass cell and allowed to evaporate under standard conditions, provided membranes with an average pore diameter of 700 mμ. By addition of small amounts of water to this " parent " solution, membranes with an average pore size as large as 2,000 mμ. could be obtained; by the addition of a good solvent of collodion such as ethylene glycol mono-ethyl ether, membranes with pore diameters down to 10 mμ. could be prepared. The pore size of this series of filters, which he called gradocol membranes, was very uniform, the ratio of the largest pore diameter to the average pore diameter being of the order of 2. It was possible to determine very accurately the average pore diameter of a batch of membranes by measuring the rate of flow of water through a square centimetre of a sample piece.

With these gradocol membranes, Elford and his colleagues investigated the sizes of viruses systematically, and a very complete account of their studies may be found in the first volume of the *Handbuch der Virusforschung* (Elford, 1938). To discover the size of a virus, the filtration " end-point " is first determined, that is, the average pore diameter of the membrane which, under optimal conditions, will just retain all virus particles assayed by an infectivity titration. In the calibration of the membranes it is necessary to assume Poiseuille's law, which is concerned with the rate of flow of water through capillaries. In view of the high water content of the membranes and their thickness, it is possible to defend this assumption: it can be shown that the channels through the membrane must be closely packed long capillaries approximately at right angles to the surfaces. But this implies that the filters cannot be regarded as simple sieves and that adsorption may be expected, even under the most favourable conditions, to play a decisive role in determining the filtration end-point. Thus it was necessary to find some independent method to enable the average pore diameter of the limiting membrane to be related to the diameter of the particle retained. Fortunately two methods were available. The larger viruses with a diameter of not less than 250 mμ. can be seen with the ordinary light microscope after they have been suitably stained. Stained particles, however, are not very satisfactory for determination of size, since the elementary bodies are enlarged to an unknown extent by the stain deposited upon

them. Virus of this size can also be seen unstained with dark-ground illumination and the resolution can be improved about twofold by using a quartz optical system, illumination from the 275 mμ. line of cadmium, and photographic recording. By this method J. E. Barnard (see Elford, 1938) was able to measure the diameter of some of the larger viruses such as vaccinia and ectromelia. The sizes of these and smaller viruses could also be estimated by centrifugation techniques. Viruses obtainable in a sufficient concentration can be accurately studied in the ultracentrifuge by the measurement of the speed of a moving boundary, disclosed by some optical method. In order to translate this information into an estimate of particle size it is necessary to assume that Stokes' law is applicable; unfortunately it is also necessary to know, with considerable accuracy, the density of the particle sedimented, and measurement of the density of viruses is very difficult. Moreover, if the virus is in too low a concentration to provide a distinct moving boundary, the sedimentation must be inferred from infectivity titrations before and after centrifugation. Owing to the vibration which occurs when a centrifuge slows down, the fluid column in ordinary tubes is greatly disturbed and under these conditions Stokes' law cannot be applied. Bechhold & Schlesinger (1931) sought to overcome this difficulty by using flat-bottomed centrifuge tubes with a disk of filter paper at the bottom to prevent redispersion of particles sedimented. They assumed that, in the thin fluid layer immediately above the filter disk, Stokes' law was applicable during centrifugation and that immediately a virus particle struck the filter paper it became enmeshed and was removed from the system. Although, on slowing down, the fluid contents of the centrifuge tube became uniformly mixed, the amount of virus lost during centrifugation could be determined by titration of the fluid.

Later Elford (1936) modified this technique. He retained the filter paper disk at the bottom of the centrifuge cup, but the virus suspension for titration was contained in an inverted capillary tube; the small diameter of this capillary tube prevented disturbance of the fluid column when the centrifuge slowed down, and in consequence Stokes' law was applicable to particles within the capillary. Infectivity titrations of the virus suspension before and after centrifugation could then be simply translated into estimates of particle size. Thus, by the independent methods of ultra-violet photography and ultracentrifugation, Elford was able to show that a limiting membrane with an average pore diameter of 1,000 mμ. retained a virus particle of the same size. But membranes with an average pore diameter of 500 mμ. retained particles only three-quarters of this pore size, while membranes of pore diameter 100 mμ. retained particles one-half of this size, and membranes of 10 mμ. retained particles one-third of this size.

With these techniques it was demonstrated that viruses which infect animal cells are indeed of different sizes, the largest having diameters of about 250 mμ., and the smallest diameters of about 10 mμ. It is a tribute to the care with which these estimates of the size of viruses were made, that when a decade or so later direct measurement of particle sizes by electron microscopy was possible, the results served only to confirm the earlier indirect estimates.

Size determination by filtration is still the method of choice when a new virus with unknown properties is investigated, and the unsatisfactory electron microscope pictures of the smallest viruses have compelled continued reliance on the indirect estimates in these cases.

Collodion gel filters are now obtainable commercially in Great Britain, Europe and the USA, and it seems probable that since they have become readily available they will be very extensively used in the immediate future. Some recent advances are reviewed by Goetz (1953).

b. Shape

Filtration methods measure the smallest diameter of a virus particle, so that by this technique alone it would not have been possible to decide whether viruses which infect animal cells are approximately spherical or whether they are rod-shaped. Estimates of size by centrifugation usually assume a spherical shape, and photographs of the larger viruses confirmed this assumption. Electron micrographs have, more recently, proved that in general the animal viruses are approximately spherical or oval. However, an apparent exception is a recently isolated strain of influenza virus. It has been shown (Mosley & Wyckoff, 1946; Chu, Dawson & Elford, 1949) that, in addition to the spherical elementary bodies which are found in the egg-adapted strains of influenza virus, suspensions of newly isolated strains also contain many long filaments. The exact nature of these filaments is in dispute, but it seems significant that their width is the same as the diameter of the spherical elementary body. Perhaps the filaments are chains of elementary bodies or alternatively the filaments may be multinucleated cells which will later fragment into elementary bodies. Owing to their characteristic shape these filaments can be readily seen by ordinary dark-ground microscopy.

c. Structure

Electron microscopy has not yet yielded any information about the structure of animal viruses in any way as unexpected as the discovery that some bacteriophages possess tails. However, it is probable that the application of this technique to virus studies has only begun to yield fruit, and the next few years should see a rich harvest of facts about morphology and virus growth cycles. In addition, the electron microscope will be used increasingly in quantitative studies as a tool for direct counts of elementary bodies.

Virus particles normally possess a rather low electron density and thus, in order to display their form, the technique of metal shadowing has been widely used (Müller, 1942; Williams & Wyckoff, 1944). However, this technique masks the details of internal structure and what is really required is some technique comparable with the staining methods of light microscopy. Perhaps multiple colour printing in register of photographs taken at different beam voltages may provide a valuable technique (Donovan & Stafford, 1952).

Dawson & McFarlane (1948) have argued that the structure of the vaccinia virus is rather against the view that viruses are small bacteria which have become intracellular parasites, but there is certainly no support for the idea that the animal viruses are large molecules.

There is an excellent review of the techniques of electron microscopy, edited by Drummond (1950).

d. Chemical Composition

In cases where adequate study has been possible it has been found that animal viruses contain protein, carbohydrate,

fat, deoxyribonucleic acid and ribonucleic acid. In a few instances more detailed chemical analyses have also been made (Knight, 1947, 1948). However, the characterization of large protein molecules and, more especially, of lipoproteins, is in such a primitive state at the present time that rapid progress in this field is unlikely.

But the chemical structure of the surface of virus elementary bodies can be studied indirectly with the help of the techniques of antigenic analysis that have been developed for the study of bacteria and large molecules. All the usual serological methods may be used, but direct immunity tests and neutralization techniques are particularly valuable as they do not require concentrated virus suspensions.

Of the in-vitro techniques the most useful is the method of complement fixation, for although it is relatively insensitive, it does not demand purified virus suspensions; moreover, this technique often possesses certain advantages over neutralization methods in establishing antigenic similarity among viruses, because the neutralization test is too specific and thus magnifies small differences (Fulton, 1952). A method of carrying out the complement-fixation test in drops on plastic sheets has been described (Fulton & Dumbell, 1949, and Plate I, fig. A). Although the results with this technique are no different from those obtainable by conventional methods using test-tubes, the ease with which the three primary reagents in the test can be independently varied, and the small volumes used, make this plate method of value in some virus problems (Kraft & Melnick, 1950; Svedmyr, Enders & Holloway, 1952; Le Bouvier, 1953).

Some authorities (Knight, 1944; Wilson Smith, 1952) have suggested that the structure of a virus elementary body depends, in part, on the host in which it is grown; they argue that the virus builds into its own body gross fragments of host material. However, the evidence to support this speculation is wholly based on the antigenic analysis of impure virus suspensions, and the experimental findings can be as well, and more simply, interpreted to mean that the purest virus suspensions as yet obtained still contain a large admixture of host material.

2. Mechanism of Adsorption to the Cell

One of the most important and least expected discoveries of the past decade was the observation that some viruses are adsorbed on to the erythrocytes of certain animals and birds, and cause a visible agglutination of the cells (Hirst, 1941; McClelland & Hare, 1941). Apart from its great practical importance for the assay of virus, for the titration of virus-neutralizing antibodies and for purification and concentration, the phenomenon of haemagglutination has been exhaustively analysed as a model for the first stage of the entry of a virus particle into a susceptible cell (Burnet, 1951). The virus particle is adsorbed specifically on to the surface of a suitable cell, probably because the surface configuration of the elementary body possesses a pattern complementary to that of the surface receptor on the cell. Under natural conditions, the virus, having come to grips with the cell, enters it in some mysterious way; in the erythrocyte model, entry does not occur and instead the virus after a time disengages itself and in so doing destroys the cellular receptor. An ingenious haemagglutination technique has been devised by Himmelweit (1949).

Although, so far, only a few viruses have been shown to agglutinate erythrocytes, the recent demonstration by Sabin & Buescher (1950), that the Japanese B virus can be persuaded to agglutinate fowl and sheep cells, raises the hope that in time the phenomenon will have general application; for to demonstrate haemagglutination by Japanese B virus many very precise conditions must be fulfilled. It is possible that the range of the haemagglutination technique may be extended by a suitable modification of the surface of the erythrocytes in the way that has proved successful for the analysis of protein antigens (Boyden, 1951; Meynell, unpublished work).

3. Enumeration of Elementary Bodies

Virus elementary bodies may be counted by a number of indirect methods, but all give estimates only of the relative numbers of bodies in a virus suspension. Thus, elementary bodies may be assayed by an infectivity titration, by haem-agglutination and by complement fixation.

Recently a direct particle-counting method has become available (Williams & Backus, 1949). The virus suspension is mixed with a suspension of polystyrene latex spheres whose absolute numerical concentration is known. The mixture is then sprayed on electron microscope specimen grids. The minute drops which fall on these grids may later be shadowed with metal and photographed in the electron microscope (Plate II, fig. D); the absolute number of virus particles in the original suspension may then be determined from the ratio of particles to polystyrene latex spheres in a number of drops.

An independent particle-counting method has been suggested by Dawson & Elford (1949). In this technique, which is applicable only to certain viruses, the elementary bodies are adsorbed to the surface of lysed chick erythrocytes. The virus-coated cells are then spread on specimen grids, shadowed with metal and photographed in the electron microscope (Plate II, fig. E).

The two methods of counting provide similar estimates (Donald, Isaacs & Fulton, in preparation).

A third technique for the enumeration of virus elementary bodies has been developed by Sharp (1949). In this method the virus is deposited on to the specimen grids by centrifugation.

It has been calculated (Friedewald & Pickels, 1944) that ten influenza virus particles are equivalent to one infectivity unit. Similarly one haemagglutinin unit is equivalent to about 10^7 virus particles, and one complement-fixing unit to about 10^8 virus particles.

4. Purification

Purified elementary body suspensions are required for serology, chemical analysis and vaccine production. Centrifugation is often used but this method yields only crudely purified preparations. Those viruses that are adsorbed to red cells and elute spontaneously may be partially purified in this way, and it is probable that the principle of the method can be generalized by the use of ion-exchange resins instead of erythrocytes (Muller & Rose, 1952).

It seems likely that in the future viruses will be purified on a preparative scale by chromatographic and phase-distribution methods. For analytical work, agar-diffusion techniques, paper chromatography and paper electrophoresis are attractive.

5. Storage

Some viruses are very hardy and will survive for long periods in aqueous suspension at + 4° C., others are more delicate and require more subtle methods of storage. The handling of viruses in the laboratory has been revolutionized by the introduction of low-temperature storage (Turner, 1938). Even delicate viruses will survive for months or years if aqueous suspensions are kept frozen at − 76° C. Craigie (1949) has argued that the critical temperature for optimal survival depends upon the freezing temperature of the eutectic solution of the salts present in biological systems, and it may be concluded that the maximum temperature permissible is about −50° C. Mechanical refrigerators which will maintain this temperature are expensive and often unreliable. Because of this, low-temperature storage of virus is most conveniently achieved with an insulated box filled with solid carbon dioxide.

Viruses may also be preserved by vacuum drying from the frozen state (Flosdorf & Mudd, 1935). However, this technique, though useful for the transport of viruses and the preservation of reserve stocks, is not so valuable as low-temperature storage for viruses in routine use, for drying kills a large proportion of the elementary bodies.

6. Multiplication of Virus inside the Cell

Cells infected with some viruses develop inclusion bodies which can be stained in many different ways; the inclusion bodies are found in the cytoplasm or the nucleus of the cell. Although some of the inclusion bodies which have been described may have been due to non-specific degenerative processes, it is probable that the majority are nests of virus elementary bodies.

Many viruses, however, do not give rise to inclusion bodies and in these cases either a suitable staining technique has not been discovered or the virus elementary bodies are distributed too diffusely to be seen.

Current ideas as to how viruses multiply are greatly influenced by the more advanced knowledge of the bacteriophage growth cycle. The author has argued elsewhere (Fulton, 1953) that it is still possible to maintain that the hereditary elements of the animal viruses are replicated in just the same way as are the chromosomes of animal and vegetable cells: thus in essence the method of reproduction of viruses is the same as that of cells which overtly multiply by binary fission.

There is a good prospect that the examination in the electron microscope of very thin sections or single layers of virus-infected cells (see Flewett & Challice, 1951) may provide decisive evidence to confirm or refute these speculations (see Plate III, fig. F).

In studies of the multiplication of bacteriophage, great use has recently been made of radioactive and stable isotope markers (Hershey & Chase, 1952); these techniques have hardly yet been applied to viruses.

7. Methods of Cultivation

Because animal viruses can multiply only within suitable animal cells the techniques used for the cultivation of viruses in the laboratory bear little resemblance to the techniques of bacteriology.

The obvious approach is to grow the virus by infecting a convenient laboratory animal. Adaptation to a new host depends on the selection, usually by chance, of a suitable member of a genotypically heterogeneous virus population. In some cases the elementary bodies of the dominant clone in the natural environment also multiply in the cells of the selected laboratory animal; in these cases primary isolations are easy. More commonly, primary isolation depends on the selection of a relatively rare mutant and the ease of isolation will depend both on the frequency of mutation and on the size of the virus population obtainable from natural sources.

For any quantitative work, and for the mass growth of viruses, cultivation by animal inoculation has many disadvantages; moreover, the use of live animals is repugnant if other means can be devised.

For these reasons, tissue culture methods provide a most attractive alternative. The chick embryo is a natural form of tissue culture which has been extensively employed since 1931 (Woodruff & Goodpasture, 1931). The techniques are very simple and are described in the monograph by Beveridge & Burnet (1946), and more recently by Karnofsky (1952). Chick embryos have been used for primary virus isolations, for assay of virus and titration of neutralizing antibodies, as a source of virus for antigens and vaccines, and in many other ways in the more academic study of viruses.

For certain purposes it is more convenient to use the de-embryonated-egg technique of Bernkopf (1949), and Fulton & Armitage (1951) have described a simple technique for the cultivation of some viruses in pieces of excised chorio-allantoic or amniotic membrane (Plate I, fig. C). This method is really a modification of the suspended cell type of culture described by Maitland & Maitland (1928).

Not all viruses have been adapted to grow in the chick embryo, though the recent success of Roca-Garcia, Moyer & Cox (1952) in adapting the MEF1 strain of poliomyelitis virus has given new hope that eventually the remaining viruses will be persuaded to multiply in avian cells.

Until a few years ago the more difficult techniques for the cultivation of mammalian cells in flasks and roller tubes (Plate I, fig. B) were sadly neglected. The history of the application of these methods to the study of viruses has been recently reviewed by Robbins & Enders (1950). The present enormous vogue in these techniques is, of course, a direct consequence of the discovery by Weller, Robbins & Enders (1949) that the poliomyelitis virus would multiply in cultures of human and monkey tissues. Poliomyelitis virus of all three types may be grown in this way and its presence in the culture may be detected by a most striking cytopathogenic effect.[1] Several other viruses are capable of producing this effect in cultures of suitable cells (see Pereira, 1953, and Plate III, fig. G). For the isolation of poliomyelitis virus these new tissue culture techniques are rapidly displacing the use of live monkeys, and the facility with which neutralizing antibodies can be titrated provides assurance of a spectacular increase in our knowledge of the epidemiology of this disease. The more obvious developments which may be expected in these tissue culture methods are modifications to permit the use of viruses that do not produce a gross cytopathogenic effect, the mass growth of virus for vaccine production, and perhaps the isolation of avirulent variants suitable for use as living vaccines.

If virus recombination in fact occurs (Burnet & Lind, 1951;

[1] See Rhodes, p. 196 of this Bulletin.

Burnet, Fraser & Lind, 1953), it is perhaps possible to dream of the artificial creation of new virus strains adapted to the needs of man. Then the mutant range of a virus could be extended to fit into a convenient scheme of cultivation, and it would no longer be necessary to await the chance emergence of avirulent strains for vaccine production.

8. Conclusions

To the biologist, higher mathematics are terrifying obstacles, but often, in the waste land of statistics made easy without mathematics, he must have wished that someone would set down a perfectly explicit text which would lead him to the heart of the matter; a more profound understanding of the principles underlying the stochastic view-point would certainly result in a great improvement in the efficient design of experiments and in the interpretation of results.

It is often imagined that virus techniques are impenetrable mysteries hidden from ordinary men. With the development of simple methods for growing viruses in tissue culture, primary isolation of many viruses may soon become a task well within the competence of any hospital or public health laboratory. The preparation of virus suspensions for serology will, no doubt, always be difficult, but there is no reason why they should not be readily available commercially, thus transferring, as in colour photography, the technical burden of production on to the shoulders of the manufacturer.

A start has been made with influenza virus antigens, which are now available through the World Health Organization, and a wider range of antigens can be purchased in the USA. However, the range of antigens required will not be the same in all parts of the world. There is little doubt that, with the progressive conquest of bacteriological diseases by chemotherapy, the study of virus diseases will become, in the future, of increasing importance. But it is well to remember that ". . . what is seen depends upon how everybody is doing everything" (Gertrude Stein, 1926).

References

Bechhold, H. (1907) *Z. phys. Chem.* **60,** 257
Bechhold, H. & Schlesinger, M. (1931) *Biochem. Z.* **236,** 387
Bernkopf, H. (1949) *Proc. Soc. exp. Biol., N.Y.* **72,** 680
Beveridge, W. I. B. & Burnet, F. M. (1946) *Spec. Rep. Ser. med. Res. Coun., Lond.* No. 256
Boyden, S. V. (1951) *J. exp. Med.* **93,** 107
Burnet, F. M. (1951) *Proc. roy. Soc.* B, **138,** 47
Burnet, F. M., Fraser, K. B. & Lind, P. E. (1953) *Nature, Lond.* **171,** 163
Burnet, F. M. & Lind, P. E. (1951) *J. gen. Microbiol.* **5,** 59 ; 67
Chu, C. M., Dawson, I. M. & Elford, W. J. (1949) *Lancet,* **1,** 602
Craigie, J. (1949) *Brit. J. Cancer,* **3,** 268
Dawson, I. M. & Elford, W. J. (1949) *Nature, Lond.* **163,** 63
Dawson, I. M. & McFarlane, A. S. (1948) *Nature, Lond.* **161,** 464
Donovan, G. E. & Stafford, E. (1952) *Nature, Lond.* **170,** 1126
Drummond, D. G. (1950) *J. R. micr. Soc.* 3rd ser. **70,** 1
Elford, W. J. (1936) *Brit. J. exp. Path.* **17,** 399
Elford, W. J. (1938) In : Doerr, R. & Hallauer, C., ed. *Handbuch der Virusforschung,* vol. 1, p. 126. Springer, Vienna
Flewett, T. H. & Challice, C. E. (1951) *J. gen. Microbiol.* **5,** 279
Flosdorf, E. W. & Mudd, S. (1935) *J. Immunol.* **29,** 389
Friedewald, W. F. & Pickels, E. G. (1944) *J. exp. Med.* **79,** 301
Fulton, F. (1952) *J. Hyg., Camb.* **50,** 265
Fulton, F. (1953) In : Fildes, P. & Heyningen, W. E. van, ed. *The nature of virus multiplication. The Second Symposium of the Society for General Microbiology held at Oxford University, April 1952,* p. 261. Cambridge University Press, London
Fulton, F. & Armitage, P. (1951) *J. Hyg., Camb.* **49,** 247
Fulton, F. & Dumbell, K. R. (1949) *J. gen. Microbiol.* **3,** 97
Fulton, F. & Isaacs, A. (1953) *J. gen. Microbiol.* **9,** 119
Goetz, W. (1953) *Amer. J. publ. Hlth,* **43,** 150
Hershey, A. D. & Chase, M. (1952) *J. gen. Physiol.* **36,** 39
Himmelweit, F. (1949) *Nature, Lond.* **164,** 650
Hirst, G. K. (1941) *Science,* **94,** 22
Karnofsky, D. A., ed. (1952) Series of papers in: *Ann. N.Y. Acad. Sci.* **55,** pages 37–344
Knight, C. A. (1944) *J. exp. Med.* **80,** 83
Knight, C. A. (1947) *J. exp. Med.* **85,** 99
Knight, C. A. (1948) *J. exp. Med.* **86,** 125
Kraft, L. M. & Melnick, J. L. (1950) *J. exp. Med.* **92,** 483
Le Bouvier, G. L. (1953) *Brit. J. exp. Path.* **34,** 300
McClelland, L. & Hare, R. (1941) *Canad. publ. Hlth J.* **32,** 530
Maitland, H. B. & Maitland, M. C. (1928) *Lancet,* **2,** 596
Mosley, V. M. & Wyckoff, R. W. G. (1946) *Nature, Lond.* **157,** 263
Müller, H. O. (1942) *Kolloidzschr.* **99,** 6
Muller, R. H. & Rose, H. M. (1952) *Proc. Soc. exp. Biol., N.Y.* **80,** 27
Pereira, H. G. (1953) *J. Path. Bact.* **65,** 259
Robbins, F. C. & Enders, J. F. (1950) *Amer. J. med. Sci.* **220,** 316
Roca-Garcia, M., Moyer, A. W. & Cox, H. R. (1952) *Proc. Soc. exp. Biol., N.Y.* **81,** 519
Sabin, A. B. & Buescher, E. L. (1950) *Proc. Soc. exp. Biol., N.Y.* **74,** 222
Sharp, D. G. (1949) *Proc. Soc. exp. Biol., N.Y.* **70,** 54
Smith, W. (1952) *Lancet,* **1,** 885
Stein, G. (1926) *Composition as explanation.* Hogarth Press, London
Svedmyr, A., Enders, J. F. & Holloway, A. (1952) *Proc. Soc. exp. Biol., N.Y.* **79,** 296
Turner, T. B. (1938) *J. exp. Med.* **67,** 61
Weller, T. H., Robbins, F. C. & Enders, J. F. (1949) *Proc. Soc. exp. Biol., N.Y.* **72,** 153
Williams, R. C. & Backus, R. C. (1949) *J. Amer. chem. Soc.* **71,** 4052
Williams, R. C. & Wyckoff, R. W. G. (1944) *J. appl. Phys.* **15,** 712
Woodruff, A. M. & Goodpasture, E. W. (1931) *Amer. J. Path.* **7,** 209
Zsigmondy, R. & Bachmann, W. (1918) *Z. anorg. Chem.* **103,** 119

TECHNIQUES FOR THE STUDY OF VIRUSES

(FIGS. A—C)

Forrest Fulton

page 186

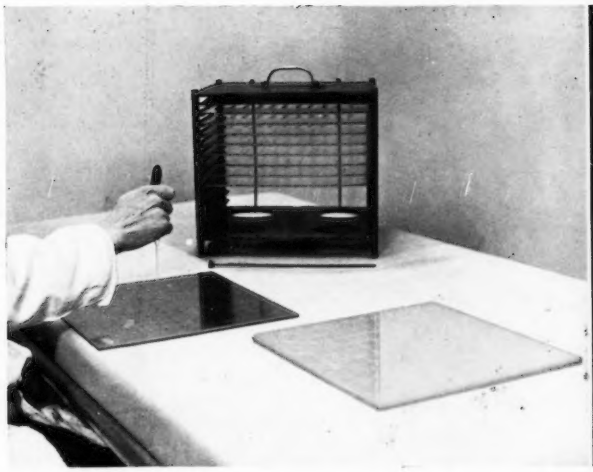


FIG. A. Apparatus for a complement-fixation technique. (For explanation of technique, see Fulton & Dumbell, 1949.)

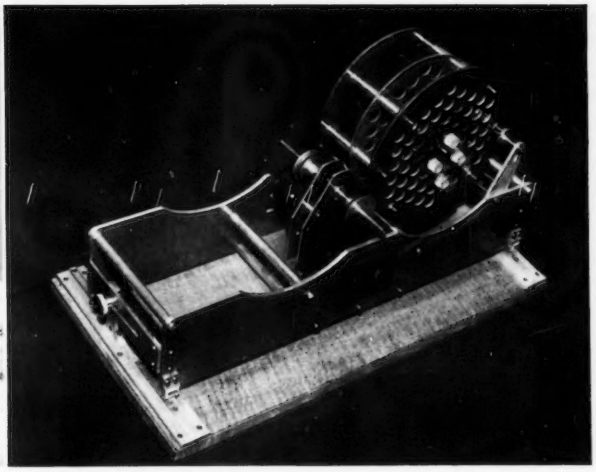
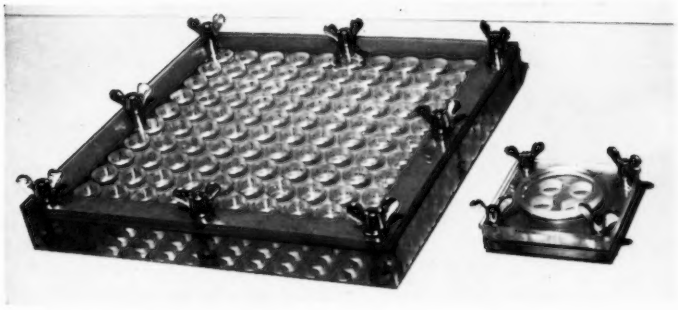


FIG. B. Apparatus for the roller-tube technique of tissue culture. Only one drum is shown. (Made by H. D. Dawson, London School of Hygiene and Tropical Medicine.) (length of baseboard: 63 cm.)

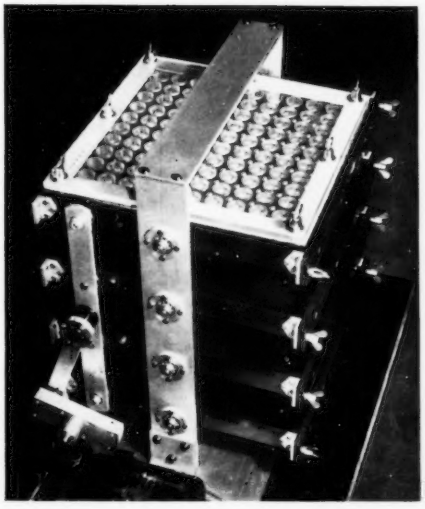
FIG. C (a, b, c). Apparatus for a method of cultivating viruses of the mumps-influenza group in surviving tissue suspensions. (For explanation of technique see Fulton & Armitage, 1951; Fulton, 1952.)

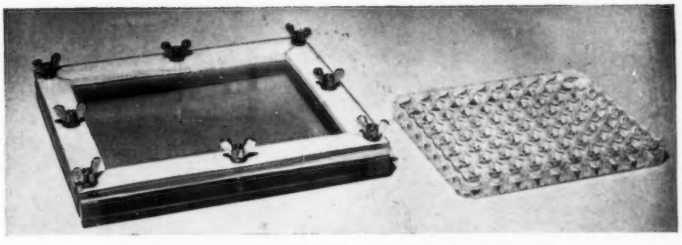


a:

Left: Perspex tray (with lid) drilled with 100 cups for assay of virus and of virus-neutralizing antibodies (dimensions of tray: 26.5 × 26.5 cm.)

Right: Perspex unit for cultivation of influenza virus on one side of the chick chorio-allantoic membrane (see Fulton & Isaacs, 1953)





b : modification of a Perspex tray available commercially

Right: Perspex tray designed for haemagglutinin assays (Prestware Ltd.)
(dimensions of tray: 18 × 14.5 cm.)

Left: Perspex box (with lid) to enclose tray and make it suitable for surviving tissue suspensions. (Box designed by J. E. Friend, London School of Hygiene and Tropical Medicine.)

c : rocking motor for 4 Perspex trays. (Made by H. D. Dawson, London School of Hygiene and Tropical Medicine.)

Electron micrographs showing methods of enumeration of influenza virus elementary bodies (Donald, Isaacs & Fulton, in preparation)

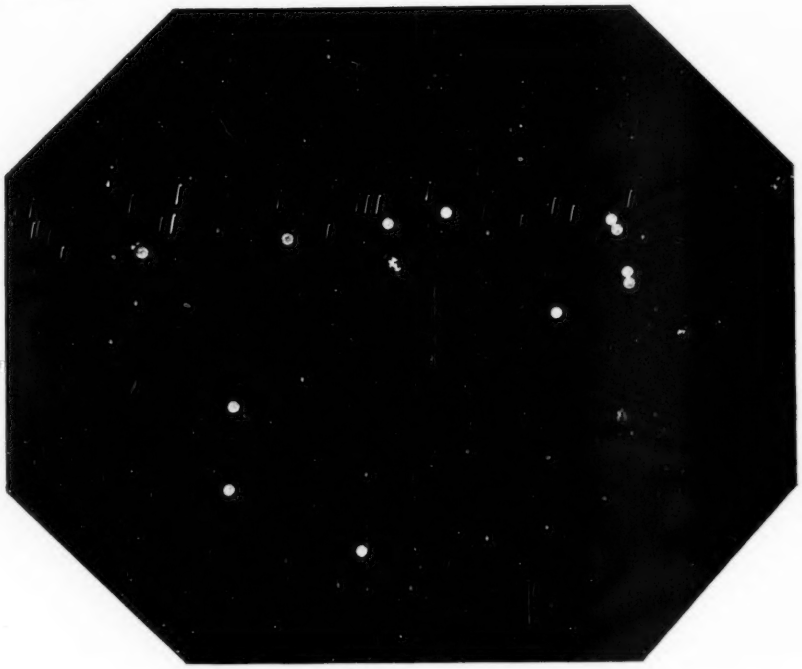


FIG. D. Droplet containing Lee strain of influenza B virus and polystyrene latex
(Technique of Williams & Backus, 1949)
The white spheres are latex particles; the very much smaller particles are virus elementary bodies
Gold manganin shadowed
(magnification × ~10,000)

FIG. E. PR8 strain of influenza A virus adsorbed on avian erythrocyte
(Technique of Dawson & Elford, 1949)
The large white central body is the nucleus of the erythrocyte; the small white spheres are virus elementary bodies
Gold manganin shadowed
(magnification × ~10,000)

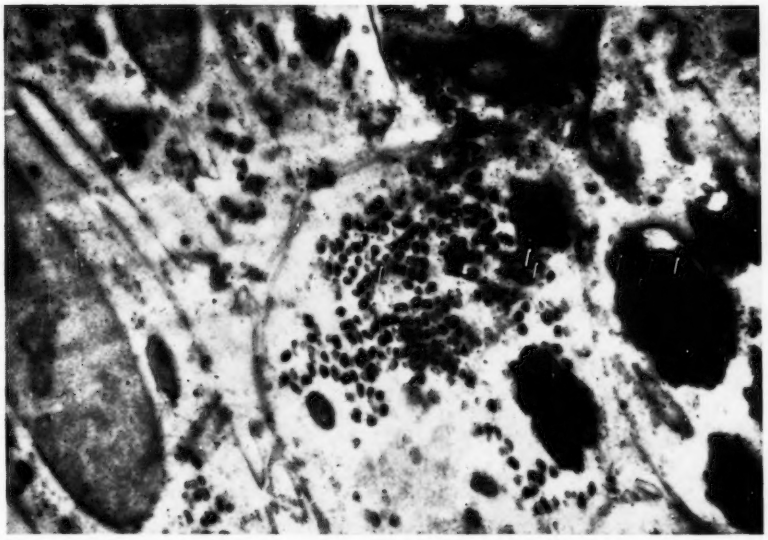
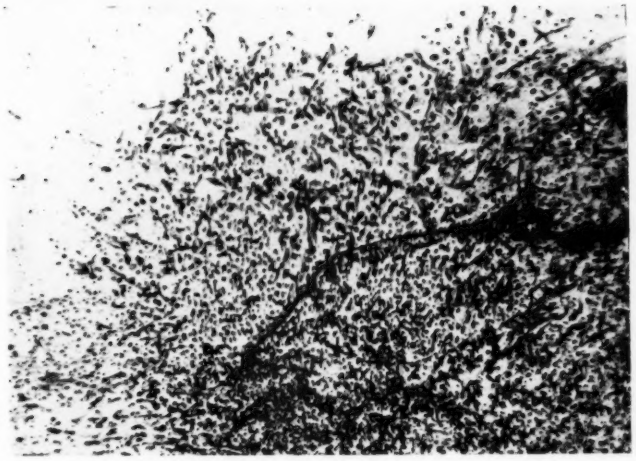
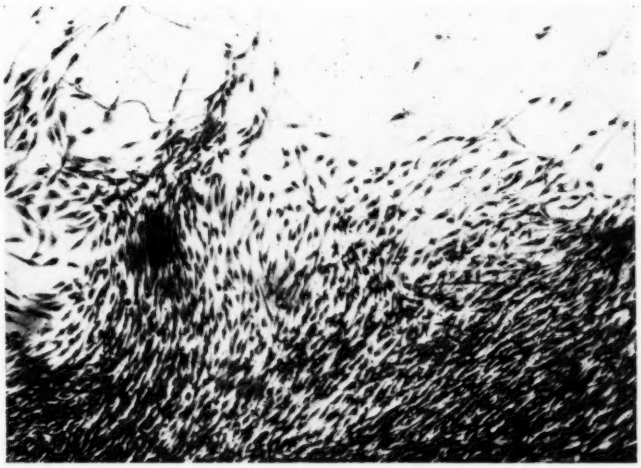


FIG. F. Electron micrograph showing vaccinia virus growing in chorio-allantoic membrane of chick embryo
Harvested after 48 hours
Fixation: buffered osmium tetroxide
Imbedding: methacrylate
Sectioned at 0.1 μ.
(magnification × 11,000)
(Preparation and electron micrograph by T. H. Flewett and G. Eaves, School of Medicine, Leeds)

FIG. G. The cytopathogenic effect in tissue culture (Preparations and photomicrographs by H. G. Pereira and A. E. H. Gompels, Common Cold Research Unit, Salisbury) (Pereira & Gompels, unpublished work)

a: culture of chick heart, 3 days after inoculation with fowl plague virus
(magnification × 78)





b: uninoculated culture of chick heart, shown for comparison with inoculated culture
(magnification × 78)

SMALLPOX AND RELATED VIRUS INFECTIONS IN MAN

(FIGS. H—M)

A. W. Downie & A. Macdonald

page 191

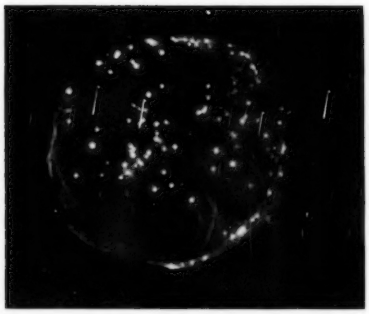


FIG. H. Lesions produced by variola virus three days after inoculation on the chick chorio-allantoic membrane
(natural size)

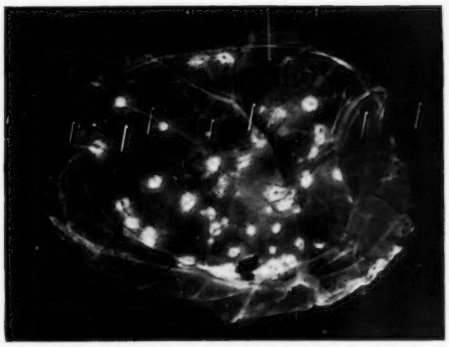


FIG. J. Lesions produced by vaccinia virus three days after inoculation on the chick chorio-allantoic membrane
(natural size)

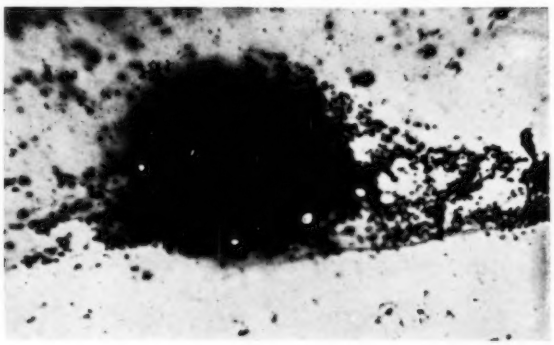


FIG. K. Smear from base of smallpox vesicle showing virus particles within the cell cytoplasm
(Stain : Gutstein ; magnification × 1,400)

FIG. L. Generalized vaccinia, showing distribution of eruption 11 days after vaccination

FIG. M. Cowpox lesions on hand of farm worker infected during milking of cows

SMALLPOX AND RELATED VIRUS INFECTIONS IN MAN

A. W. DOWNIE M.D. D.Sc.
A. MACDONALD M.A. M.D.

Department of Bacteriology
University of Liverpool

1 Variola major and variola minor
  a Pathogenesis
  b The virus
  c Laboratory tests in diagnosis
2 The spread of smallpox
3 Prevention and control
4 Specific prophylaxis and treatment by passive immunization
5 Cowpox in man
References

Smallpox is still endemic in large areas of the world and as long as this is so no country can be assured of permanent freedom from the infection. From the main reservoirs of infection, in India, south-east Asia, central Africa and Central America, travellers, by ship or plane, may introduce the infection into countries where, because of improved standards of living, vaccination and other measures of control, the disease has ceased to be endemic. In Great Britain the disease has been reintroduced from abroad on many recent occasions and the infecting case has frequently been a partially protected, vaccinated individual. Consequently the illness in these first cases is often mild and atypical and has sometimes not been immediately recognized as smallpox. Outbreaks which have followed such reintroductions have been limited by the vigorous methods taken to control spread but the mortality has been high (20–30%) and unvaccinated individuals have suffered severely. Indeed recent experience has shown that the major form of the disease is no less virulent than in the eighteenth century when it was one of the main endemic scourges. Alastrim or the minor form of variola, on the other hand, as shown by the outbreak in Lancashire in 1952, has also remained true to type as a mild clinical illness with negligible mortality.

The outbreaks of variola that have occurred in Britain during the past twelve years have offered favourable opportunities for epidemiological study, since the source of infection has usually been traced and the outbreaks have been small. Moreover, developments in laboratory methods of diagnosis have now made such tests more precise and reliable than they are in most other virus infections. Consequently, atypical cases can be recognized with certainty and the laboratory can help in the retrospective identification of missed cases, thus allowing the spread of infection to be followed more completely.

This article is mainly, but not exclusively, concerned with experience gained in recent outbreaks of smallpox in Britain and particularly with those aspects of the disease which come within the range of the laboratory investigator. In addition there is a brief description of human infections with the virus of cowpox.

1. Variola Major and Variola Minor

On clinical and epidemiological evidence there is no doubt that the two varieties of smallpox can readily be distinguished during outbreaks. In the last 15 years, with the exception of 1952, smallpox in Britain has been of the major type with the severity and mortality associated with that disease. During 1952, however, there was an outbreak of variola minor involving over 130 cases, almost all of them characteristically mild infections and similar to the cases seen during the extensive endemic period from 1923 to 1935. There were no deaths. In an analysis of 13,686 cases personally studied between 1928 and 1934, Marsden (1948) has compared and contrasted the clinical picture of variola minor with that described by Ricketts in his classical study of variola major (Ricketts & Byles, 1908). Marsden points out that modified or mild smallpox may occur in two circumstances: either the infecting virus is of low virulence as in variola minor or the patient has increased resistance as in variola major in persons vaccinated. In modified smallpox of both types the lesions of the focal rash, when compared with those of severe smallpox, are " small, superficial, and of rapid development, hurrying through their course, or aborting and assuming a granulomatous termination " (Marsden, 1948). However, one fundamental difference must be noted. Variola minor remains a mild disease when transferred to vaccinated or unvaccinated contacts, whereas modified or mild variola major may be, and often is, the source of the most severe type of the infection. The incubation period and pre-eruptive period of illness are of the same duration in both major and minor forms of the disease, although secondary fever is less common in the latter. From an analysis of smallpox in previously vaccinated persons there is some evidence that vaccinial immunity may be more durable against the minor than against the major form of the disease (Marsden, 1948).

a. Pathogenesis

In older studies, such as those of Councilman, Magrath & Brinckerhoff (1904), bacterial infection was thought to play an important part in the pathology of the disease. Recent work has tended to minimize the effect of bacteria in determining the clinical and pathological picture. A recent report on the morbid anatomy of smallpox by Bras (1952) was based on the findings in 177 fatal cases. The majority of the patients had received antibiotic treatment and there was no histological evidence that any of the pathological features described were due to secondary bacterial infection. Nor do the negative bacteriological findings of blood culture in confluent and haemorrhagic cases and the failure to culture bacteria from skin pustules suggest that bacteria play an essential role. Antibiotics do not influence the early toxaemic phase, and this also supports the view that the severity of smallpox and the mortality from it are due to the virus and are in most cases independent of secondary bacterial infection (Downie, 1950, 1953).

The virus of smallpox probably enters the tissues through the mucous membrane in the upper respiratory tract and

during the incubation period may multiply in lymphoid and other tissues. The patient is not infectious at this time so that it is unlikely that there is an open lesion on the mucosal surface. No evidence is available from laboratory investigations on patients during this period but experimental studies on the natural pox disease of mice (Fenner, 1948) suggest that primary multiplication takes place in the lymphatic tissue draining the site of entry. This may be followed by a transient viraemia whereby reticulo-endothelial cells throughout the body become infected; proliferation of virus in these cells leads to the secondary and more intense viraemia which ushers in the clinical illness (Downie, 1951a). Observations on the transmission of infection from mothers with smallpox to the foetus in utero (Marsden & Greenfield, 1934; Dixon, 1948) suggest that infection of the foetus from the maternal circulation occurs about the time of onset of clinical illness in the mother. The skin lesions must result from dissemination of virus by the blood, and indeed virus can be isolated from blood during the early stages of smallpox in a proportion of patients (Downie, McCarthy & Macdonald, 1950; MacCallum, McPherson & Johnstone, 1950). This viraemia, however, is probably short-lived, for virus is not usually to be found in the blood after the second day of fever except in cases which are to prove fatal. The improvement in the clinical condition that occurs in the average patient with the development of the skin eruption is probably due to the formation of antibody which can be detected in the blood soon afterwards. Secondary fever, which is associated with pustulation of the skin lesions, is presumably caused by the absorption of the products of cell necrosis in the absence of secondary bacterial infection.

b. The Virus

Variola virus resembles other viruses of the pox group in size, being 200–300 mμ. in diameter and brick-shaped when seen in dried films in the electron microscope. In infected cells it gives rise to the same type of granular acidophilic cytoplasmic inclusions as does vaccinia; these are almost certainly agglomerations of virus particles. Variola virus differs from vaccinia in that its host range is limited to man and monkeys and it cannot be propagated in series in ordinary laboratory animals. Fortunately it can be grown readily on the chorio-allantois of chick embryos. In this tissue it produces characteristic lesions different from those of vaccinia (Plate IV, figs. H, J). On continued subculture on the chick embryo it retains these characters for many generations; the lesions show no tendency to change towards those of vaccinia. Propagation on the chorio-allantois has permitted more intensive study of the virus in the laboratory.

The antigenic structure of variola virus is similar to that of vaccinia. Both release soluble antigen in infected tissue and the presence of this antigen in the skin lesions provides a means for serological diagnosis by complement-fixation tests. Both viruses produce a haemagglutinin (Nagler, 1944; North, 1944), which is distinct from soluble antigen (Chu, 1948) and which may also be detected by in-vitro tests. Cross-immunity tests in monkeys with the viruses of variola and vaccinia have shown that the two are closely related immunologically (Gordon, 1925; Horgan & Haseeb, 1939). Recent work on the serological and immunological relationships between the viruses of variola major, variola minor, vaccinia and cowpox has shown that the viruses of variola major and minor are indistinguishable and that between these and the other two only minor differences can be elicited (Downie & McCarthy, 1950; Downie & Macdonald, 1950).

c. Laboratory Tests in Diagnosis

In typical cases of smallpox, laboratory confirmation of the diagnosis is rarely required. In clinically doubtful cases the tests likely to give assistance at various stages of illness are shown in Table I. In the smallpox case of moderate severity, virus can be detected in the blood in only a proportion of cases and then only in small amount during the first 24 hours of illness. In the fulminating haemorrhagic case, however, virus can usually be found in the blood and may be detectable up to the time of death. In some of these cases antigen may be found in the blood serum by complement-fixation tests (Downie, McCarthy, Macdonald, MacCallum & Macrae, 1953). Virus can be isolated from the skin and typical changes may be found on histological examination of skin removed post mortem.

In Tables II and III there are set out the results of diagnostic tests carried out in our laboratory on material from cases of variola major (Table II) and variola minor (Table III). The results of examinations of blood for virus and antigen are not included in these tables. The negative microscopic findings on stained smears from skin lesions in confirmed smallpox cases were due in part to the unsuitable nature of the material submitted. Smears carefully prepared from lesions in the papular and vesicular stage and properly stained give a high percentage of positive results (see Plate IV, fig. K). The

TABLE I. LABORATORY TESTS IN DIAGNOSIS OF SMALLPOX

Stage of illness	Material to be submitted	Microscopic examination of smears from skin lesions	Culture on chick embryo chorio-allantois	Detection of antigen by complement-fixation test	Detection of antibody by complement-fixation test
Pre-eruptive illness	Blood		±	±	−
Macular and papular stage	Smears (on slides) from skin lesions	+	+	±	−
Vesicular .	Vesicle fluid	+	+	+	
	Smears (on slides) from base of vesicles	+	+	+	−
Pustular .	Pustule fluid	±	+	+	
	Blood				±
Crusting stage	Crusts	−	+	+	
	Blood				+
Later . .	Blood				+
Time required for completion of test		30 min.	2–3 days	24 hr.	24 hr.

\+ test is usually positive.
± test may or may not be positive.
− test is usually negative.

TABLE II. RESULTS OF LABORATORY TESTS IN VARIOLA MAJOR*

	Tests on material from skin lesions			Complement fixation for antibody in blood
	Microscopic examination of smears	Isolation of virus on chorio-allantois of chick embryo	Complement fixation for antigen	
Positive . .	34	61	35	54
Negative . .	13	2	1	3
Total number of specimens . .	47	63	36	57

* These specimens were obtained from 114 cases of variola major, in all of which the diagnosis was confirmed by one or other of these tests.

method of detecting virus on the chorio-allantois of 12–14-day-old chick embryos is perhaps the most reliable laboratory test (Torres & Teixeira, 1935; Buddingh, 1938; Bohls & Irons, 1942; Downie & Dumbell 1947a). This is the simplest laboratory method of distinguishing cases of smallpox from generalized vaccinia (Plate IV, fig. L) occurring in vaccinated persons during an epidemic. The lesions produced by variola virus are quite different in appearance from those due to vaccinia virus on this tissue (see Plate IV, figs. H, J). The complement-fixation test for antigen is a sensitive and reliable method if sufficient material is collected from skin lesions. The test for antibody in the serum is likely to be positive by complement fixation or antihaemagglutinin titrations after the first week of illness but is of no value in recently vaccinated individuals. While we agree with Collier, Smit & von Herde (1950) that the titre of antibody is usually higher in smallpox patients than it is after vaccination we have not found the differences sufficiently great or constant to make the test of diagnostic value in recently vaccinated persons. Antibody is not usually to be found in the serum by the complement-fixation technique more than 12 months after vaccination. Consequently in cases of variola *sine eruptione* with a history of vaccination more than 12 months previously this test may be of assistance in diagnosis.

We have not been able to distinguish by laboratory tests between cases of variola major and minor. Torres (1935) has described cytological differences in the skin lesions of these two conditions, but we have not found such differences on microscopic examination of the lesions on the chick embryo chorio-allantois.

Atypical cases of variola may be mistaken clinically for chickenpox, erythema multiforme, pustular acne, drug rashes and other non-specific illnesses associated with a skin eruption. In none of these conditions do we find the microscopic picture presented by suitably stained smears from smallpox lesions and in none will the test for specific antigen or culture test on the chick embryo be positive. In our experience cases of generalized vaccinia are not as uncommon as the literature suggests. We have during the last few years examined material from several cases in adults vaccinated for the first time during smallpox outbreaks. In such cases the focal lesions in the skin may be similar to those of variola although the distribution is not so decidedly centrifugal as in that disease (Plate IV, fig. L). As noted above, these cases can be readily distinguished by the character of the pocks produced on the chorio-allantois by virus from the skin lesions. In the rare cases of generalized infection with herpes simplex virus, lesions are produced on the chorio-allantois by fluid from the vesicles, but these can be distinguished from those of smallpox by their smaller size, their histology and by serological means (Kipping & Downie, 1948).

2. The Spread of Smallpox

Although usually considered to be an extremely infectious disease, smallpox does not have the epidemic potential of influenza or even of chickenpox. In the records of recent outbreaks, spread has usually been traced to direct contact between cases. It is true that spread may sometimes be indirect, for example, through infected bed linen, as in the infection of laundry workers (Stallybrass, 1931) and in the recent Brighton outbreak (Cramb, 1951; Parker, 1952). The virus may survive for months or even years outside the body (Downie & Dumbell, 1947b) and outbreaks in Lancashire have been traced to imported cotton (Corbin, 1915; Ministry of Health, 1934).

Epidemiological evidence suggests that most cases of smallpox are not infectious from the onset of illness but become so at the time of appearance of the skin eruption. This view was strongly held by Ricketts and is supported by recent observations. For example, in the Glasgow outbreak of variola major in 1949, individuals were infected from the first undiagnosed case and many of them were not detected and isolated until the appearance of the rash; yet none of these gave rise to further clinical cases although vaccination of contacts had not been carried out at this time (Laidlaw & Horne, 1950). At the time the eruption occurs there is a large amount of virus in the lower layer of the skin epidermis, but, because of the impermeable nature of this tissue, virus is not shed into the environment until the rupture of vesicles and pustules at a later stage. In the mouth and pharynx, however, this protective layer is absent and lesions break down early, and thus rapidly render the saliva infective. We have been unable to demonstrate any inhibitory effect on variola virus of saliva from unvaccinated or recently vaccinated persons; and our attempts to demonstrate variola virus in the mouth

TABLE III. RESULTS OF LABORATORY TESTS IN VARIOLA MINOR*

	Tests on material from skin lesions			Complement fixation for antibody in blood
	Microscopic examination of smears	Isolation of virus on chorio-allantois of chick embryo	Complement fixation for antigen	
Positive . .	23	73	26	18
Negative . .	17	2	5	4
Total number of specimens . .	40	75	31	22

* These results were obtained on 89 confirmed cases of variola minor.

during the first two days of the pre-eruptive fever have so far been unsuccessful. In acute cases of purpura variolosa, however, where bleeding may occur under the skin and from mucous membranes soon after the onset of illness, the patient may, as suggested by Stallybrass, be infectious throughout the febrile period (Stallybrass, 1931). These considerations have obvious epidemiological importance, and during an outbreak the diagnosis of smallpox should always be borne in mind for any acutely ill patient with haemorrhages in the skin or mucous surfaces.

3. Prevention and Control

Since the introduction of the National Health Service in Britain, compulsory vaccination has been abolished and the over-all vaccination rate among infants has now fallen to less than 20% (Conybeare, 1950). Even before this official change of policy, probably less than 50% of the population had been vaccinated against smallpox. These figures do not, however, mean that faith in the power of vaccination to protect against smallpox has been lost; rather that in present circumstances compulsory vaccination is unnecessary because smallpox is no longer an endemic disease in England and because it is felt that methods short of universal vaccination will serve to protect the country against severe epidemics. The implementation of sanitary regulations at ports has helped to prevent the introduction of smallpox into this island with its world-wide commercial contacts. The speed of modern air travel, however, makes the introduction of smallpox more difficult to check, and the importation of infection in such materials as cotton is liable to evade the most rigorous measures of control at seaports and airports. When the disease appears in Britain, control depends on the prompt detection and isolation of cases and the tracing, listing, vaccination and surveillance of all contacts. In 1946 fifteen separate introductions of variola major into England resulted in only forty secondary cases, although such good fortune has not always attended vigorous public health control measures. Local mass vaccination, such as was practised in New York in 1946, is considered unnecessary so long as the disease occurs only in the direct line of contact or an association with known cases can be reasonably established (Bradley, 1947). It is noteworthy that the only recent outbreak of variola minor—in 1952—was considerably greater in extent than any post-war experience of variola major. In variola minor, owing to the mild nature of the illness, more patients escape detection because they are often ambulant while still infectious. Should circumstances change and the danger of greater prevalence of smallpox arise, the present measures of control may have to be revised.

4. Specific Prophylaxis and Treatment by Passive Immunization

While the antibiotics may be of value in preventing secondary infection in the pustular stage of smallpox, there is no evidence that they have any effect on the early toxic phase of the illness. Nor do the scanty records in the literature suggest that even large doses of convalescent serum, given after the illness has declared itself, will prevent a fatal issue in the most severe cases of variola major. However, preparations of concentrated antibody in the form of γ-globulin from the serum of smallpox convalescents or recently vaccinated individuals have not yet been adequately tested in smallpox. The success of antibody preparations against a few other virus infections would seem to call for similar investigation in variola major. It is probable that concentrated antibody given before the onset of illness might serve to inactivate some of the virus during the stage of viraemia and so lessen the severity of the subsequent attack on the skin and other tissues. As is well known, convalescent serum added to a suspension of variola virus renders the virus incapable of infecting the chorio-allantois of the chick embryo. Immune γ-globulin might therefore be given with advantage during the incubation period to unvaccinated contacts of variola major patients. This method of prophylaxis by passive immunization would certainly seem worthy of trial.

5. Cowpox in Man

This disease is of minor importance but is mentioned briefly because of its historical interest and the relationship of the causal virus to that of variola. The disease appears to be endemic in Holland (Dekking, 1950) and in England, and is probably more prevalent than the reported outbreaks suggest (Donaldson & Waldron, 1951; Downie, 1951b). The disease in cattle is milder than the pock diseases of other animals. The lesions are confined to the teats and udders, and infection in man is contracted by direct contact during milking. The lesions in milkers usually occur on the hands (Plate IV, fig. M), particularly between the forefinger and thumb, but sometimes on the forearm or face. In unvaccinated persons the disease is more severe than in the vaccinated; the local lesion is often associated with lymphangitis, lymphadenitis and some fever and general upset. Very rarely there is a generalized eruption and one fatal case with meningo-encephalitis has been recorded (Verlinde, 1951). The vesicular lesions in man, as in cattle, resemble those produced by vaccination, although there is more congestion and the serous exudate may be blood-stained. Healing of the lesions may take several weeks.

The cow is the normal host of this virus and in Great Britain infection of cattle from human beings recently vaccinated (i.e., with vaccinia virus) has not been established, although outbreaks elsewhere have been attributed to such a source (Hamburger, 1948). The strains of virus isolated in Great Britain from lesions in cattle, or in milkers infected from them, have been identical and differ constantly in certain respects from the strains of vaccinia virus used for the preparation of vaccine lymph. All cowpox strains produce typical deep red haemorrhagic lesions on the chick embryo chorio-allantois. When discrete they retain their haemorrhagic nature, a finding which serves readily to distinguish them from vaccinia strains in current use. Histologically the lesions on the chorio-allantois of the chick embryo, and on the skin of guinea-pigs, rabbits and humans, show typically large, compact, strongly acidophilic inclusions quite unlike the inclusions produced by vaccinia virus. By these properties cowpox viruses can be identified. The strains of vaccinia virus that have long been propagated in lymph establishments are of uncertain origin. They may have been derived originally from cowpox strains but, if so, their superficial characters have changed during continued artificial passage in the epidermis of domestic and laboratory animals. Immunologically and in host range, cowpox virus closely resembles vaccinia virus and its immunological relationship to variola is sufficiently close to justify Jenner's belief in the immunity to smallpox of those who had previously had the cowpox.

REFERENCES

Bohls, S. W. & Irons, J. V. (1942) *Amer. J. publ. Hlth*, **32**, 300
Bradley, W. H. (1947) *J. R. sanit. Inst.* **67**, 559
Bras, G. (1952) *Docum. Med. geogr. trop.* **4**, 303
Buddingh, G. J. (1938) *Amer. J. Hyg.* **28**, 130
Chu, C. M. (1948) *J. Hyg., Camb.* **46**, 49
Collier, W. A., Smit, A. M. & Herde, A. F. von (1950) *Z. Hyg. InfektKr.* **131**, 557
Conybeare, E. T. (1950) *Mon. Bull. Minist. Hlth Lab. Serv.* **9**, 120
Corbin, H. E. (1915) *Proc. R. Soc. Med.* **8**, pt 2, 89
Councilman, W. T., Magrath, G. B. & Brinckerhoff, W. R. (1904) *J. med. Res.* **11**, 12
Cramb, R. (1951) *Publ. Hlth, Lond.* **64**, 123
Dekking, F. (1950) *Tijdschr. Diergeneesk.* **75**, 248
Dixon, C. W. (1948) *J. Hyg., Camb.* **46**, 351
Donaldson, E. & Waldron, D. H. (1951) *Mon. Bull. Minist. Hlth Lab. Serv.* **10**, 40
Downie, A. W. (1950) In: Bedson, S. E., Downie, A. W., MacCallum, F. O. & Stuart-Harris, C. H. *Virus and rickettsial diseases*, p. 151. Arnold, London
Downie, A. W. (1951a) *Brit. med. J.* **2**, 251
Downie, A. W. (1951b) *Lancet*, **1**, 419
Downie, A. W. (1953) In: *History of the Second World War. Medicine and pathology*, edited by Cope, H. Z., p. 497. Her Majesty's Stationery Office, London
Downie, A. W. & Dumbell, K. R. (1947a) *J. Path. Bact.* **59**, 189
Downie, A. W. & Dumbell, K. R. (1947b) *Lancet*, **1**, 550
Downie, A. W. & McCarthy, K. (1950) *Brit. J. exp. Path.* **31**, 789
Downie, A. W., McCarthy, K. & Macdonald, A. (1950) *Lancet*, **2**, 513
Downie, A. W., McCarthy, K., Macdonald, A., MacCallum, F. O. & Macrae, A. D. (1953) *Lancet*, **2**, 164
Downie, A. W. & Macdonald, A. (1950) *J. Path. Bact.* **62**, 389
Fenner, F. (1948) *J. Path. Bact.* **60**, 529
Gordon, M. H. (1925) *Spec. Rep. Ser. med. Res. Coun., Lond.* No. 98
Hamburger, R. J. (1948) *Ned. Tijdschr. Geneesk.* **92**, 1161; *Bull. Hyg., Lond.* **23**, 566 [Abstract]
Horgan, E. S. & Haseeb, M. A. (1939) *J. Hyg., Camb.* **39**, 615
Kipping, R. H. & Downie, A. W. (1948) *Brit. med. J.* **1**, 247
Laidlaw, S. I. A. & Horne, W. A. (1950) *Med. Offr*, **83**, 187
MacCallum, F. O., McPherson, C. A. & Johnstone, D. F. (1950) *Lancet*, **2**, 514
Marsden, J. P. (1948) *Bull. Hyg., Lond.* **23**, 735
Marsden, J. P. & Greenfield, C. R. M. (1934) *Arch. Dis. Childh.* **9**, 309
Ministry of Health (1934) *On the state of the public health: annual report of the Chief Medical Officer of the Ministry of Health for the year 1933*, p. 34. His Majesty's Stationery Office, London
Nagler, F. P. O. (1944) *Aust. J. exp. Biol. med. Sci.* **22**, 29
North, E. A. (1944) *Aust. J. exp. Biol. med. Sci.* **22**, 105
Parker, W. S. (1952) *J. R. sanit. Inst.* **72**, 105
Ricketts, T. F. & Byles, J. B. (1908) *Diagnosis of smallpox.* Cassell, London
Stallybrass, C. O. (1931) *The principles of epidemiology and the process of infection*, pages 329 & 338. Routledge, London
Torres, C. M. (1935) *Proc. R. Soc. Med.* **29**, 1525
Torres, C. M. & Teixeira, J. de C. (1935) *C. R. Soc. Biol., Paris*, **118**, 1023
Verlinde, J. D. (1951) *Tijdschr. Diergeneesk.* **76**, 334

RECENT ADVANCES IN THE STUDY OF POLIOMYELITIS, WITH PARTICULAR REFERENCE TO PREVENTION

A. J. RHODES M.D. F.R.C.P.

Professor of Virus Infections
School of Hygiene, University of Toronto

Director of Research
The Hospital for Sick Children
Toronto, Canada

1 The pathogenesis of poliomyelitis
  a Experimental aspects
  b Human aspects
2 The origin of poliomyelitis outbreaks
3 The tissue cultivation of poliomyelitis viruses
4 Antigenic typing of poliomyelitis viruses
5 Prospects of vaccination against human poliomyelitis
6 γ-Globulin in the prevention of poliomyelitis
7 Conclusions
References

Poliomyelitis continues to attract more attention in the literature than any other virus disease, and recently several additions to our knowledge have been made. It is the object of this review to present to the general medical reader the more significant of these, with particular reference to the all-important problem of prevention by immunization, active or passive. Space does not allow of extensive citation of publications, but more extensive bibliographies are readily available (van Rooyen & Rhodes, 1948; National Foundation for Infantile Paralysis, 1951a; International Poliomyelitis Congress, 1952; Rivers, 1952; Rhodes & van Rooyen, 1953).

1. The Pathogenesis of Poliomyelitis

a. Experimental Aspects

Poliomyelitis virus behaves as a neurotrope in the monkey, and the classical studies of Fairbrother & Hurst (1930), and those of Howe & Bodian (1942) have shown that after intracerebral, intraneural, or intranasal inoculation, virus spreads by axonal pathways. Monkeys can also be infected by mouth, by intradermal and intramuscular injection, and intravenously. An alternative route of transportation by lymph and blood has therefore been postulated. Of particular interest are two recent series of observations. Faber, Silverberg & Dong (1951, 1953) reported that, if cynomolgus monkeys were inoculated by rubbing virus on to the buccal and oropharyngeal mucosa, or by feeding, virus rapidly reached the peripheral ganglia by nerve fibres. Histological lesions were found in ganglia within 48 hours, and viraemia was detected contemporaneously. Infection might remain localized in the ganglia or ascend to the central nervous system. In the other series it was found that virus circulated in the blood of chimpanzees and cynomolgus monkeys infected by feeding; this occurred in the incubation period and was transient (Horstmann, 1952; Bodian, 1952a). It is not yet apparent whether viraemia is an essential stage in the dissemination of virus in monkeys infected peripherally, or whether it represents a "spill-over" from a focus in regional ganglia. However, the demonstration that low levels of antibody will protect from paralysis monkeys inoculated by peripheral routes lends support to the former suggestion (Bodian, 1951, 1952b).

The concept that virus proliferates outside the nervous system receives support from another source, for inoculated monkeys and chimpanzees develop serum antibody without histological evidence of nervous involvement. It is probable that antibody develops as a sequel to multiplication in the alimentary mucosa or the viscera.

Hormonal mechanisms are also concerned in pathogenesis, for susceptibility to infection is increased by treatment with cortisone. Shwartzman has made such observations in hamsters and has detected viraemia (Shwartzman, 1950, 1952; Shwartzman & Fisher, 1952). A similar increased susceptibility has been observed in mice (Findlay & Howard, 1952), and monkeys (Ainslie, Francis & Brown, 1951; Foster, Sigel, Henle & Stokes, 1951).

b. Human Aspects

Some light on pathogenesis is thrown by consideration of the clinical course of individual illnesses (Horstmann, 1949; Russell, 1949, 1952; Faber, 1950; Bodian, 1952a). Thus, in children the "dromedary" or two-phase illness is common. The first phase illness begins with fever, headache, sore throat, anorexia, vomiting, and listlessness. After a remission of a few days, the second phase illness begins with fever, headache, pain, stiffness, irritability and drowsiness; paralysis usually follows. The explanation of the "dromedary" illness is obscure, but it is probable that the first phase represents a reaction to viral proliferation in viscera, regional ganglia, or even central nervous system. During the remission, virus disseminates more widely, prior to multiplying actively in the central nervous system, with the eventual development of clinical manifestations.

Additional evidence comes from neurohistological examinations of the peripheral and central nervous systems. Faber (1950) concludes that the virus enters through the mucosa of the pharynx or mouth and thence invades through peripheral nerves. He envisages the following sequence: deposition of virus on mucosa of nose, mouth, and intestine; entry into peripheral nerve-endings; ascent to peripheral ganglia and eventually to the central nervous system. The virus is primarily neurotropic, and is "locked away" inside the nervous system from shortly after penetration of the superficial mucosae. If this is correct, it would appear that the progress of the virus cannot be halted except by some agent acting on the surface mucosa or on susceptible neurones.

An alternative theory has recently been revised by Bodian (1952a), who postulates that virus multiplies first in the alimentary mucosa, then in the organs associated with the blood, and finally in the central nervous system and sensory ganglia. Viraemia is initiated by the escape of virus from the primary alimentary multiplication sites and is brought to an end by the appearance of antibody; although virus has been found in the blood on only two occasions (Ward, Horstmann & Melnick, 1946; Koprowski, Norton & McDermott, 1947), tests carried out early in the incubation should be positive.

Whatever mechanisms are concerned, serological studies make it clear that infection is only occasionally manifested by nervous illness (Turner, Hollander, Buckley, Kokko & Winsor, 1950). Several factors may convert such inapparent infections into paralytic illnesses. These precipitating factors operate at various stages after deposition of virus on the mucosae. For example, a pharynx recently traumatized by tonsillectomy may afford a ready means of access to nerve fibres, and unknown factors may likewise increase the permeability of the pharyngeal and intestinal mucosae. Later, when invasion has occurred and virus circulates in the blood, virus may be localized in some traumatized area, proliferate, and proceed by nerve fibres to the central nervous system, with the production of paralysis of the traumatized area. Attention has recently been drawn to the role of a particular form of trauma, for intramuscular injections of alum-precipitated diphtheria toxoid and pertussis vaccine may be followed within four weeks by paralysis of the inoculation site (see McCloskey, 1950, 1951; Hill & Knowelden, 1950; and reviews by Faber, 1951, and by Rhodes, 1953). It is probable that, in these persons, virus settled in the inoculation site, proliferated, and then passed to the central nervous system by nerve fibres. Support for this concept is afforded by observations that less irritant antigens injected subcutaneously do not appear to increase the hazard of developing paralysis (Rhodes, 1953; see Bousfield, 1950, 1951, 1952).

Two further aspects deserve mention. Thus, pregnancy apparently serves to increase susceptibility to paralysis, probably through hormonal mechanisms (see Garlick, 1948; Anderson, Anderson, Skaar & Sandler, 1952). Finally, there seems to be an increased incidence of paralysis in those who continue physical exertion after the onset of the major illness, as if exercise favours the multiplication of virus present in, but not fully disseminated through, the central nervous system (Russell, 1947, 1949, 1952; Hargreaves, 1948; Horstmann, 1950; Albrecht & Locke, 1951; Sabin, 1949; Brahdy & Katz, 1951).

It is evident that further work is needed but, in the meantime, it is suggested that viraemia is an essential stage in the dissemination of poliomyelitis virus through the human body. Whether virus in the blood penetrates the central nervous system directly, or whether it reaches this system by passage along nerve fibres, is not known.

2. The Origin of Poliomyelitis Outbreaks

There is now agreement that poliomyelitis is essentially spread by human contact, and that stools form the main source of virus. Epidemiological inquiry usually reveals a history of contact with a previous case, often a child suffering from an abortive illness. Exceptionally, food or drink may be contaminated by a human excreter and serve to infect a group of persons. The evidence has recently been reviewed by Sabin (1951a), who concludes as follows: human faeces are the most significant source of virus; the period of communicability may extend for longer than a few days after onset; the mouth is the usual portal of entry; faecal-feeding flies may act as carriers of virus; and it is likely that several methods of transmission are involved.

The major unsolved problem in epidemiology is the seasonal incidence in North America, Europe, and other temperate zones. In the tropics, where the evidence is increasing that poliomyelitis is endemic, cases occur the year round. Various attempts have been made to link seasonal incidence with meteorological conditions, and the most recent are those of Armstrong (1950, 1951; Armstrong, Dixon & Chadwick, 1952). Cold weather itself does not necessarily interfere with the transmission of the virus; not only may summer outbreaks continue into the winter, but epidemics beginning in this season have been recorded. In the epidemic that involved Canadian Eskimos in the winter of 1948-49, the temperature was −30° F.[1] (Adamson, Moody, Peart, Smillie, Wilt & Wood, 1949; Peart, 1949; Rhodes, Clark, Goodfellow & Donohue, 1949; Clark & Rhodes, 1951, 1952; Wood, Clark, Shimada & Rhodes, 1953). Another Canadian winter outbreak occurred in an isolated settlement near the Quebec-Labrador boundary (Pavilanis & Frappier, 1951).

In the absence of any evidence for extra-human survival, it has been postulated that in the temperate zones virus survives between epidemics in the human population of urban areas. As paralytic cases are very rare during the winter, it is necessary to suggest that virus passes sporadically from person to person, causing a chain of abortive illnesses or carrier states. With the onset of warmer weather, for some reason there is increased transfer of virus, and thus an increased incidence of infection.

In a number of outbreaks in isolated communities, for instance, the epidemics on St Helena and in the Canadian north country, the epidemiological evidence has pointed to the introduction of infection by an "outsider". Similar introductions must take place in more settled areas also. The state of immunity of the population would be expected to play a major role in determining the outcome of such an introduction of infection.

The concept of varying degrees of "invasiveness" of virus has been advanced as an explanation of the varying intensity of infection in different outbreaks (see Burnet, 1952; Sabin, 1951b). "Invasiveness" or "virulence" is not, however, a property that can be measured in the laboratory, and strains isolated from fatal cases may fail to induce severe disease in monkeys. "Invasiveness" appears to depend at least as much on the susceptibility of the host as on any property of the virus.

3. The Tissue Cultivation of Poliomyelitis Viruses

A most important discovery was made by Enders and his associates Robbins and Weller, at the Children's Medical Center, Boston, Massachusetts, when it was found that all three types of virus proliferate in tissue cultures of human and monkey organs (Enders, Weller & Robbins, 1949; Weller, Robbins & Enders, 1949; Robbins, Enders & Weller, 1950; Robbins, Enders, Weller & Florentino, 1951; Enders, 1952; Robbins, Weller & Enders, 1952; Weller, Enders, Robbins & Stoddard, 1952). This work has been extended by others (e.g., Smith, Chambers & Evans, 1951; Ledinko, Riordan & Melnick, 1952; Riordan, Ledinko & Melnick, 1952; Franklin, Duncan, Wood & Rhodes, 1952; Salk, Lewis, Bennett & Youngner, 1952; Scherer & Syverton, 1952; Wood, Franklin, Clark, Duncan & Rhodes, 1952b; Youngner, Lewis, Ward & Salk, 1952).

In the technique originally described, fragments of minced human embryonic organs, monkey testis or kidney are placed in Erlenmeyer flasks and covered with nutrient (Hanks-Simms)

[1] t° F. = ⅝ (t − 32)° C.—Ed.

fluid. The tissues metabolize rapidly; after 24 hours the nutrient is changed and virus added. Further replacements of nutrient are made every 3 to 4 days. The tissues metabolize for 28 days or more, and actively liberate virus into the nutrient fluid. Virus growth interferes with tissue metabolism, which ceases appreciably before metabolism in uninfected cultures. This technique is of value in the isolation of viruses, and for the production of virus-infected fluid.

In the second technique, fragments of human embryonic skin and muscle, uterus, monkey testis or kidney are embedded in a plasma clot in a test-tube. A nutrient containing serum and embryo extract is added and the tubes are rotated horizontally; there are several advantages in the use of a synthetic nutrient elaborated in Toronto and known as " Mixture 199 " (Morgan, Morton & Parker, 1950; Franklin *et al.* 1952; Wood *et al.* 1952b). Within 4 to 7 days, an outgrowth of " fibroblasts " develops round each explant; with kidney, the cells are epithelial. The nutrient is then replaced and virus added. Virus enters the cells and rapidly produces a " cytopathogenic " change seen microscopically in the form of disintegration and necrosis; this change is inhibited specifically by antiserum.

Tissue culture has been very widely used, and the main applications include: isolation of viruses from specimens; titration of viruses and antisera; typing of viruses; and preparation of virus-containing fluid for use as antigen in complement-fixation tests (Svedmyr, Enders & Holloway, 1952) or as a vaccine.

4. Antigenic Typing of Poliomyelitis Viruses

Earlier studies (e.g., Bodian, Morgan & Howe, 1949; Morgan, 1949a; Kessel & Pait, 1950) showed the existence of three types. These observations were extended by a group of USA investigators who examined 100 strains; the distribution was as follows: type 1 (Brunhilde), 85 strains; type 2 (Lansing), 12 strains; and type 3 (Leon), 3 strains (National Foundation for Infantile Paralysis, 1951b). Various methods of vaccinating and challenging monkeys were used, particularly a method in which the untyped virus was injected intramuscularly, the inoculum being incorporated with an oily adjuvant; serum from the immunized monkeys was tested for capacity to neutralize fixed amounts of the three known types (Salk, Lewis, Youngner & Bennett, 1951). More recently, typing in tissue culture has largely replaced these methods.

The majority of strains of virus have been recovered by the inoculation of monkeys or of tissue cultures. Only a few strains such as the Lansing have been adapted to rodents; these appear to belong to type 2, and are known as " Lansing-like " (Armstrong, 1939; Bodian, 1949a).

5. Prospects of Vaccination against Human Poliomyelitis

Monkeys recovered from infection are resistant to " challenge " (reinfection); this state is specific, and operates effectively only against infection with the same type, although there is some cross-resistance (Bodian, 1949c).

A number of earlier workers, especially Brodie (1932, 1934) and Kolmer (1937), demonstrated that monkeys could be rendered resistant also by the inoculation of vaccines containing live virus, or virus treated with formalin or other agents. The work was carried to a stage in the 1930's at which children were inoculated with virus treated with chemicals. Cases of paralysis were attributed to the use of these materials (Kolmer, 1936), and no further attempts to vaccinate children were made.

The subject has been placed on a more secure basis by the experiments of Morgan and her associates at Johns Hopkins University, Baltimore. This work was carried out mainly with the Lansing strain by methods more quantitative than those used formerly. This study established that in vaccinated monkeys the degree of resistance is correlated with the level of serum antibody (Morgan, 1949b).

When poliomyelitis virus inactivated by formalin is employed to vaccinate monkeys, the antibody response is poor (Morgan, 1948). This is of serious significance from the viewpoint of human use. However, antibody response is increased by incorporating virus in paraffin oil (Bayol) with an emulsifier (Arlacel). These adjuvants, usually with acid-fast bacilli, have been used by Freund in studies with bacteria (Freund & Bonanto, 1944) and with rabies virus (Lipton & Freund, 1950). Use of antigens containing nervous tissue incorporated with oil and acid-fast bacilli results in the development of " allergic " encephalomyelitis (Morgan, 1947). This complication does not arise if acid-fast bacilli are omitted. Monkeys vaccinated with adjuvant-treated virus develop higher antibody levels than those receiving an aqueous suspension (Ward, Rader, Lipton & Freund, 1950; Rhodes, Clark, Shimada, Donohue & Ritchie, 1952). This phenomenon has been extensively investigated by Salk, who has prepared antisera to many poliomyelitis strains (Salk *et al.* 1951). He has also carried out studies with influenza virus in monkeys (Salk & Laurent, 1952) and in humans (Salk, Bailey & Laurent, 1952); the latter studies emphasize the safety of the adjuvants for man.

Until recently, most of the experimental work on vaccination of monkeys was carried out with suspensions of nervous tissue. Although such suspensions treated with formalin stimulate antibodies in man (Howe, 1952), they do not form ideal sources of virus; the hazard of " allergic " encephalomyelitis is a deterrent.

It is doubtful whether the recent experience gained in the vaccination of monkeys could have been applied to man unless Enders *et al.* (1949) had discovered that virus can be produced in tissue culture. Tissue culture fluids incorporated with adjuvants and injected intramuscularly stimulate a brisk antibody response in monkeys (Salk, Lewis, Bennett & Youngner, 1952; Wood *et al.* 1952b).

There is still a prospect that poliomyelitis virus vaccines can be produced in chick embryos, and it has been reported that a Lansing-type strain (MEF1) adapted to hamsters has been passed by the allantoic and yolk-sac routes (Cabasso, Stebbins, Dutcher, Moyer & Cox, 1952; Moyer, Accorti & Cox, 1952; Roca-García, Moyer & Cox, 1952). To date, neither type 1 nor type 3 virus has been so adapted.

The possibility of feeding an adapted strain to humans by mouth has also been explored (Koprowski, Jervis & Norton, 1952). Such a procedure, even if " covered " by an accompanying injection of γ-globulin, is too hazardous to recommend for general introduction. Variants entirely devoid of virulence have not been discovered.

It will be obvious that the prospects of developing for trial an inactivated virus vaccine, probably from tissue cultures, are very good. It should then be possible to resume the field studies abandoned 20 years ago. The objective of such a

vaccination programme would be to stimulate the development of sufficient serum antibody to protect against extra-alimentary multiplication of virus and invasion of the central nervous system. Vaccination would be recommended for children between the ages of six months or a year and the early teens. Between these ages the level of serum antibody is known to be low.

Very recently, Salk (1953) has reported the results of inoculating a small group of persons with virus grown in tissue culture and made non-infectious by treatment with formaldehyde; the inoculum acted as an antigen and stimulated the development of antibody.

6. γ-Globulin in the Prevention of Poliomyelitis

Poliomyelitis-neutralizing antibody is present in γ-globulin (Enders, 1944); this fraction prepared from pooled human plasma collected during the Second World War by the American National Red Cross contains antibody to the three types (Bodian, 1949b). γ-Globulin acts as an effective prophylactic against infection in mice injected cerebrally (Rhodes & Clark, 1951), or in monkeys inoculated by the oral or intramuscular routes (Bodian, 1951, 1952b). Significant antibody levels develop in the serum of monkeys when γ-globulin is injected intramuscularly; the level then declines logarithmically, the half-life being about eight days (Bodian, 1951). In man, the half-life is somewhat longer (Dixon, Talmage, Maurer & Deichmiller, 1952). Relatively low antibody levels are sufficient to protect monkeys from developing paralysis after peripheral inoculation or feeding (Bodian, 1951, 1952b).

Encouraged by results in the prevention of measles and infectious hepatitis, and by experimental evidence, Hammon and his associates Coriell and Stokes recently conducted a field trial of human γ-globulin, prepared from large donor pools of plasma collected by the American National Red Cross, as a prophylactic (Hammon, Coriell & Stokes, 1952a, 1952b; Hammon, Coriell, Wehrle, Klimt & Stokes, 1952). These trials were conducted in 1951 and 1952 in areas in the USA where severe epidemics were in progress. Of the 54,772 children from 1 to 11 years of age inoculated, half received γ-globulin intramuscularly in a dosage similar to that used in the prevention of measles, and the other half received only gelatin. Ninety cases of paralysis subsequently occurred in the two groups. Considering the cases with onset in the first week after inoculation, there was no significant difference between the numbers in the γ-globulin group (12) and those in the gelatin group (16). However, in the second and subsequent weeks there were many fewer cases of paralysis in the γ-globulin group (14) than in the gelatin group (48). It was concluded that γ-globulin conferred a significant degree of protection for a period of about five weeks. Studies carried out in Toronto, Canada, by Wood, Clark, McKendry & Rhodes (1952a) in children and adults suggested that serum antibody levels in Hammon's series were low. Small amounts of serum antibody appear, therefore, to confer protection in man as in monkeys. It would seem that the role of γ-globulin would be to confer short-term protection at the time of a severe epidemic of paralytic poliomyelitis.

7. Conclusions

This review presents to the general medical reader a summary of some recent experimental work on poliomyelitis. Particular attention is paid to problems of immunization, both active by means of vaccines, and passive by means of γ-globulin. The subject of pathogenesis is first reviewed, and it is shown that the orthodox concept of poliomyelitis as a strictly neurotropic infection has recently been challenged. New evidence suggests that virus probably circulates in the blood for a brief period before the onset of the major illness. It would seem likely that this stage of viraemia can be prevented by the prophylactic administration of γ-globulin, and that the level of antibody needed to achieve this effect is low. The recent work on tissue cultivation of poliomyelitis viruses is then reviewed, and it is shown that this technique has many applications, particularly in the diagnosis and prevention of the disease. The prospects of developing a vaccine for human use are then discussed, and it is concluded that the most promising vaccine at the moment is one prepared from tissue cultures and treated with formalin.

The studies of Hammon on the prophylaxis of poliomyelitis by means of γ-globulin are then reviewed. It would seem that the successful results were due to the neutralization of virus in the blood stream. The future with regard to the prevention of poliomyelitis by means of vaccines or in certain circumstances by γ-globulin would appear to be very hopeful.

REFERENCES

Adamson, J. D., Moody, J. P., Peart, A. F. W., Smillie, R. A., Wilt, J. C. & Wood, W. J. (1949) *Canad. med. Ass. J.* **61,** 339

Ainslie, J. D., Francis, T., jr & Brown, G. C. (1951) *J. Lab. clin. Med.* **38,** 344

Albrecht, R. M. & Locke, F. B. (1951) *J. Amer. med. Ass.* **146,** 769

Anderson, G. W., Anderson, G., Skaar, A. & Sandler, F. (1952) *Amer. J. Hyg.* **55,** 127

Armstrong, C. (1939) *Publ. Hlth Rep., Wash.* **54,** 1719; 2302

Armstrong, C. (1950) *Amer. J. publ. Hlth,* **40,** 1296

Armstrong, C. (1951) *Amer. J. publ. Hlth,* **41,** 1231

Armstrong, C., Dixon, J. P. & Chadwick, W. L. (1952) *Amer. J. publ. Hlth,* **42,** 1246

Bodian, D. (1949a) *Proc. Soc. exp. Biol., N.Y.* **70,** 1

Bodian, D. (1949b) *Proc. Soc. exp. Biol., N.Y.* **72,** 259

Bodian, D. (1949c) *Amer. J. Hyg.* **49,** 200

Bodian, D. (1951) *Amer. J. Hyg.* **54,** 132

Bodian, D. (1952a) *Amer. J. Hyg.* **55,** 414

Bodian, D. (1952b) *Amer. J. Hyg.* **56,** 78

Bodian, D., Morgan, I. M. & Howe, H. A. (1949) *Amer. J. Hyg.* **49,** 234

Bousfield, G. (1950) *Med. Offr,* **84,** 29

Bousfield, G. (1951) *Lancet,* **1,** 1028

Bousfield, G. (1952) *Med. Offr,* **87,** 108

Brahdy, M. B. & Katz, S. H. (1951) *J. Amer. med. Ass.* **146,** 772

Brodie, M. (1932) *J. exp. Med.* **56,** 493

Brodie, M. (1934) *Science,* **79,** 594

Burnet, F. M. (1952) *Med. J. Aust.* **1,** 169

Cabasso, V. J., Stebbins, M. R., Dutcher, R. M., Moyer, A. W. & Cox, H. R. (1952) *Proc. Soc. exp. Biol., N.Y.* **81,** 525

Clark, E. M. & Rhodes, A. J. (1951) *Canad. J. med. Sci.* **29,** 216

Clark, E. M. & Rhodes, A. J. (1952) *Canad. J. med. Sci.* **30,** 390

Dixon, F. J., Talmage, D. W., Maurer, P. H. & Deichmiller, M. (1952) *J. exp. Med.* **96,** 313

Enders, J. F. (1944) *J. clin. Invest.* **23,** 510

Enders, J. F. (1952) *J. Immunol.* **69,** 639

Enders, J. F., Weller, T. H. & Robbins, F. C. (1949) *Science,* **109,** 85

Faber, H. K. (1950) *Pediatrics, Springfield,* **6,** 488

Faber, H. K. (1951) *Pediatrics, Springfield,* **7,** 300

Faber, H. K., Silverberg, R. J. & Dong, L. (1951) *J. exp. Med.* **94,** 455

Faber, H. K., Silverberg, R. J. & Dong, L. (1953) *J. exp. Med.* **97,** 69
Fairbrother, R. W. & Hurst, E. W. (1930) *J. Path. Bact.* **33,** 17
Findlay, G. M. & Howard, E. M. (1952) *J. Pharm., Lond.* **4,** 37
Foster, C., Sigel, M. M., Henle, W. & Stokes, J., jr (1951) *J. Lab. clin. Med.* **38,** 359
Franklin, A. E., Duncan, D., Wood, W. & Rhodes, A. J. (1952) *Proc. Soc. exp. Biol., N.Y.* **79,** 715
Freund, J. & Bonanto, M. V. (1944) *J. Immunol.* **48,** 325
Garlick, C. H. (1948) *N. Z. med. J.* **47,** 599
Hammon, W. M., Coriell, L. L. & Stokes, J., jr (1952a) *J. Amer. med. Ass.* **150,** 739
Hammon, W. M., Coriell, L. L. & Stokes, J., jr (1952b) *J. Amer. med. Ass.* **150,** 750
Hammon, W. M., Coriell, L. L., Wehrle, P. F., Klimt, C. R. & Stokes, J., jr (1952) *J. Amer. med. Ass.* **150,** 757
Hargreaves, E. R. (1948) *Brit. med. J.* **2,** 1021
Hill, A. B. & Knowelden, J. (1950) *Brit. med. J.* **2,** 1
Horstmann, D. M. (1949) *Amer. J. Med.* **6,** 592
Horstmann, D. M. (1950) *J. Amer. med. Ass.* **142,** 236
Horstmann, D. M. (1952) *Proc. Soc. exp. Biol., N.Y.* **79,** 417
Howe, H. A. (1952) *Amer. J. Hyg.* **56,** 265
Howe, H. A. & Bodian, D. (1942) *Neural mechanisms in poliomyelitis.* Commonwealth Fund, New York
International Poliomyelitis Congress (1952) *Poliomyelitis: papers and discussions presented at the Second International Poliomyelitis Conference.* Lippincott, London
Kessel, J. F. & Pait, C. F. (1950) *Amer. J. Hyg.* **51,** 76
Kolmer, J. A. (1936) *Amer. J. publ. Hlth,* **26,** 126
Kolmer, J. A. (1937) *J. Immunol.* **32,** 341
Koprowski, H., Jervis, G. A. & Norton, T. W. (1952) *Amer. J. Hyg.* **55,** 108
Koprowski, H., Norton, T. W. & McDermott, W. (1947) *Publ. Hlth Rep., Wash.* **62,** 1467
Ledinko, L., Riordan, J. T. & Melnick, J. L. (1952) *Amer. J. Hyg.* **55,** 323
Lipton, M. M. & Freund, J. (1950) *J. Immunol.* **64,** 297
McCloskey, B. P. (1950) *Lancet,* **1,** 659
McCloskey, B. P. (1951) *Med. J. Aust.* **1,** 613
Morgan, I. M. (1947) *J. exp. Med.* **85,** 131
Morgan, I. M. (1948) *Amer. J. Hyg.* **48,** 394
Morgan, I. M. (1949a) *Amer. J. Hyg.* **49,** 225
Morgan, I. M. (1949b) *J. Immunol.* **62,** 301
Morgan, J. F., Morton, H. J. & Parker, R. C. (1950) *Proc. Soc. exp. Biol., N.Y.* **73,** 1
Moyer, A. W., Accorti, C. & Cox, H. R. (1952) *Proc. Soc. exp. Biol., N.Y.* **81,** 513
National Foundation for Infantile Paralysis (1951a) *A bibliography of infantile paralysis, 1789–1949: with selected abstracts and annotations,* 2nd ed., edited by Fishbein, M., Salmonsen, E. M. & Hektoen, L. Lippincott, London
National Foundation for Infantile Paralysis. Committee on Typing (1951b) *Amer. J. Hyg.* **54,** 191
Pavilanis, V. & Frappier, A. (1951) *Ann. Inst. Pasteur,* **81,** 583
Peart, A. F. W. (1949) *Canad. J. publ. Hlth,* **40,** 405
Rhodes, A. J. (1953) *Canad. med. Ass. J.* **68,** 107
Rhodes, A. J. & Clark, E. M. (1951) *Proc. Soc. exp. Biol., N.Y.* **76,** 379
Rhodes, A. J., Clark, E. M., Goodfellow, A. & Donohue, W. L. (1949) *Canad. J. publ. Hlth,* **40,** 418
Rhodes, A. J., Clark, E. M., Shimada, F. T., Donohue, W. L. & Ritchie, R. C. (1952) *Canad. J. med. Sci.* **30,** 54
Rhodes, A. J. & van Rooyen, C. E. (1953) *Textbook of virology,* 2nd ed. Williams & Wilkins, Baltimore
Riordan, J. T., Ledinko, N. & Melnick, J. L. (1952) *Amer. J. Hyg.* **55,** 339
Rivers, T. M., ed. (1952) *Viral and rickettsial infections of man,* 2nd ed. Lippincott, London
Robbins, F. C., Enders, J. F. & Weller, T. H. (1950) *Proc. Soc. exp. Biol., N.Y.* **75,** 370
Robbins, F. C., Enders, J. F., Weller, T. H. & Florentino, G. L. (1951) *Amer. J. Hyg.* **54,** 286
Robbins, F. C., Weller, T. H. & Enders, J. F. (1952) *J. Immunol.* **69,** 673
Roca-Garcia, M., Moyer, A. W. & Cox, H. R. (1952) *Proc. Soc. exp. Biol., N.Y.* **81,** 519
Russell, W. R. (1947) *Brit. med. J.* **2,** 1023
Russell, W. R. (1949) *Brit. med. J.* **1,** 465
Russell, W. R. (1952) *Poliomyelitis.* Arnold, London
Sabin, A. B. (1949) *Ann. intern. Med.* **30,** 40
Sabin, A. B. (1951a) *J. Pediat.* **39,** 519
Sabin, A. B. (1951b) *Amer. J. publ. Hlth,* **41,** 1215
Salk, J. E. (1953) *J. Amer. med. Ass.* **151,** 1081
Salk, J. E., Bailey, M. L. & Laurent, A. M. (1952) *Amer. J. Hyg.* **55,** 439
Salk, J. E. & Laurent, A. M. (1952) *J. exp. Med.* **95,** 429
Salk, J. E., Lewis, L. J., Bennett, B. L. & Youngner, J. S. (1952) *Fed. Proc.* **11,** 480
Salk, J. E., Lewis, L. J., Youngner, J. S. & Bennett, B. L. (1951) *Amer. J. Hyg.* **54,** 157
Scherer, W. P. & Syverton, J. T. (1952) *J. exp. Med.* **96,** 389
Shwartzman, G. (1950) *Proc. Soc. exp. Biol., N.Y.* **75,** 835
Shwartzman, G. (1952) *Proc. Soc. exp. Biol., N.Y.* **79,** 573
Shwartzman, G. & Fisher, A. (1952) *J. exp. Med.* **95,** 347
Smith, W. M., Chambers, V. C. & Evans, C. A. (1951) *Proc. Soc. exp. Biol., N.Y.* **76,** 696
Svedmyr, A., Enders, J. F. & Holloway, A. (1952) *Proc. Soc. exp. Biol., N.Y.* **79,** 296
Turner, T. B., Hollander, D. H., Buckley, S., Kokko, U. P. & Winsor, C. P. (1950) *Amer. J. Hyg.* **52,** 323
van Rooyen, C. E. & Rhodes, A. J. (1948) *Virus diseases of man,* 2nd ed. Nelson, New York
Ward, R., Horstmann, D. M. & Melnick, J. L. (1946) *J. clin. Invest.* **25,** 284
Ward, R., Rader, D., Lipton, M. M. & Freund, J. (1950) *Proc. Soc. exp. Biol., N.Y.* **74,** 536
Weller, T. H., Enders, J. F., Robbins, F. C. & Stoddard, M. B. (1952) *J. Immunol.* **69,** 645
Weller, T. H., Robbins, F. C. & Enders, J. F. (1949) *Proc. Soc. exp. Biol., N.Y.* **72,** 153
Wood, W., Clark, E. M., McKendry, J. B. J. & Rhodes, A. J. (1952a) *Proc. Soc. exp. Biol., N.Y.* **80,** 522
Wood, W., Franklin, A. E., Clark, E. M., Duncan, D. & Rhodes, A. J. (1952b) *Proc. Soc. exp. Biol., N.Y.* **81,** 434
Wood, W., Clark, E. M., Shimada, F. T. & Rhodes, A. J. (1953) *Canad. J. med. Sci.* **31,** 207
Youngner, J. S., Lewis, L. J., Ward, E. N. & Salk, J. E. (1952) *Amer. J. Hyg.* **55,** 347

COXSACKIE VIRUSES

J. O'H. TOBIN M.A. B.M. B.Ch. Dp. Bact.

Department of Bacteriology
University of Manchester

1 General characters
2 Infections associated with C virus
  a Herpangina
  b Bornholm disease
  c Aseptic meningo-encephalitis
  d Poliomyelitis
  e Other diseases
3 Epidemiology
4 Diagnosis of Coxsackie infections
  a Virus isolation
  b Serological diagnosis
5 Summary
References

Coxsackie is the name of a Hudson river town in the USA where lived the two subjects from whom Dalldorf & Sickles (1948) isolated the first member of the Coxsackie or "C" group of viruses. Since then knowledge of these agents has increased rapidly. As they have been found in the faeces and throat washings of normal people as well as of patients suffering from many different kinds of illness, an understanding of their exact part in human disease has lagged behind the knowledge of their physical and biological properties as demonstrated in the laboratory. However, their aetiological role in certain diseases is being gradually established, especially in the two clinical conditions of herpangina and Bornholm disease.

1. General Characters

The C viruses do not comprise a homogeneous group but can be separated into a number of different agents, having the common property of high infectivity for new-born mice. Although all can induce the formation of antibodies in mice of all ages only a few are able to produce recognizable symptoms or death in any but suckling animals. Young hamsters and *mérions* (Lépine, Chaumont & Blusson, 1952) are also susceptible. Chimpanzees and cynomolgus monkeys can be infected subclinically, becoming pharyngeal and intestinal carriers of virus for short periods. Some strains have been grown in tissue culture (Slater & Syverton, 1950; Dalldorf, 1952) and in fertile eggs (Dalldorf, 1952; Peers, Ransom & Huebner, 1952).

On histological grounds these viruses can be divided broadly into two groups (Dalldorf, 1950; Gifford & Dalldorf, 1951). Those of group A are characterized by their ability to produce widespread myositis in the skeletal musculature of suckling mice without other observable lesions, together with the clinical picture of flaccid paralysis. Depending on the strain of virus, these animals are susceptible to infection until the second or third week of life. Group B viruses on the other hand, although generally capable of producing a myositis which is more focal in distribution than that of the A group, also give rise to a panniculitis, mostly affecting the fat pads between the scapulae and the sucking pads, and often to encephalitis, the animals dying with paralysis of the spastic type. Some of the B strains will also produce pancreatitis in both suckling and adult mice, but this lesion depends to a great extent on what material is used for passage (Dalldorf & Gifford, 1952). Cortisone may enhance the susceptibility of older mice to infection of the pancreas but appears to have little effect on the pathogenicity of those strains unable to produce this lesion a priori (Sulkin, Wallis & Donaldson, 1952). Occasionally the group B viruses also produce myocarditis and hepatitis in young animals. Mice of 1 to 2 days of age are most suitable for the growth of the B group.

Affected muscles show hyaline degeneration with the marked cellular infiltration of an active repair process. In the brain the group B viruses produce a widespread necrosis often involving the whole cerebrum. In the fat pads necrosis begins at the periphery of the lobules and spreads inwards, followed by regeneration of the damaged cells. Often areas of calcification can be seen in these lesions. In older animals an infected pancreas is involved in a severe inflammatory process with destruction of the acini and their replacement by fat, but leaving the islets unaffected (Pappenheimer, Kunz & Richardson, 1951). Animals with this lesion generally lose weight and may die, but often show no obvious signs of illness.

The A and B groups often differ from one another in other respects (Table I) although some strains may have properties characteristic of either group. Stanley, Dorman & Ponsford (1953) have suggested a third group for strains which produce only encephalitis in suckling mice. However, as their strains showed a serological relationship to a member of the B group, further work, will be necessary before it is decided whether this further subdivision of the C viruses is valid or not.

These viruses are very small, being of the same order of size as those of poliomyelitis and mouse encephalitis. Although gradocol filtration experiments would indicate that some particles are only 10–15 mμ. in diameter (Himmelweit,

TABLE I. GENERAL DIFFERENTIAL CHARACTERISTICS OF GROUPS A AND B COXSACKIE VIRUSES

Characteristic	Group A	Group B
Optimum age of mice for virus inoculation	1–5 days	1–2 days
Average titre (ID_{50}) of 10% carcass suspension	10^{-6}–10^{-8}	10^{-3}–10^{-6}
Incubation period in suckling mice	2–6 days	4–12 days
Type of paralysis	Flaccid	Spastic
Myositis in suckling mice	Generalized	Focal or absent
Panniculitis in suckling mice	Absent	Marked
Encephalitis in suckling mice	Absent	Often present
Pancreatitis in weaned mice	Absent	Often present
Titre of neutralizing antibody response in man (with 100 ID_{50} of virus)	High ($\frac{1}{250}$ – $\frac{1}{6250}$)	Moderate ($\frac{1}{50}$ – $\frac{1}{250}$)
Chief clinical picture produced in man	Herpangina	Bornholm disease

Findlay & Howard, 1950), electron micrographs and ultracentrifuge determinations would indicate that the majority have a diameter nearer 30–38 mμ. (Briefs, Breese, Warren & Huebner, 1952).

They are extremely stable agents, some being destroyed only after heating at 60° C. for at least 30 minutes. They are resistant to ether and to all antibiotics and are not precipitated by protamine sulphate. They survive in a wide range of pH and salt concentrations and are unaffected by 5% Lysol. They can be preserved in glycerol and at temperatures below — 20° C., but may not always survive drying from the frozen state. They are readily inactivated by 0.1 N-hydrochloric acid and formalin (Melnick, 1951).

Serologically the viruses are heterogeneous, there being at least eleven different types in group A and five in group B. Dalldorf (1952) has suggested type numbers within each group, as similar viruses have been given many different names.

2. Infections Associated with C Virus

As the first isolations of C viruses were made from patients suffering from paralytic poliomyelitis, many cases of this disease were investigated for their presence and numerous strains were found both in the paralytic and non-paralytic disease. They have also been isolated from cases of summer grippe, aseptic meningitis, herpangina, Bornholm disease and epidemic myalgia, and non-specific fever.

The two main conditions in which they appear to play a definite aetiological role are herpangina and Bornholm disease. As both of these diseases can cause a pleocytosis in the cerebrospinal fluid (c.s.f.), some cases of aseptic meningitis should also be considered as due to C virus infection.

a. Herpangina

This acute disease is characterized by fever, dysphagia, pharyngitis and small ulcerative lesions on the fauces and tonsils. It was first described by Zahorsky (1920) and is widespread during the summer months in the USA, particularly in children. It can be accompanied by vomiting, anorexia, abdominal pain and convulsions (Parrott, Ross, Burke & Rice, 1951). Huebner and his colleagues have shown that this disease is associated with many Coxsackie viruses of the A group and they found them in either the faeces or throat washings of 85% of their cases (Huebner, Beeman, Cole, Bergelman & Bell, 1952). That the viruses were the causal agent of this syndrome was confirmed also by their occurrence in contacts and in near relatives of the sick (40–60%), and only rarely in other people (3.5%). Moreover, a rise in neutralizing antibody to the strain of virus causing the disease was observed in all cases in which early and late serum specimens were obtained. Another outbreak of this disease in the USA, which involved group A viruses, has been reported recently (Kravis, Hummeler & Sigel, 1952), while in Great Britain Stuart-Harris (personal communication, 1952) has also isolated an A10 type of virus from a case of herpangina. Findlay & Howard (1950) instilled an A2 strain into the nose of a volunteer who 46 hours later had slight fever and pain below the right scapula, which lasted for about 48 hours or so. But two laboratory workers infected accidentally with an A4 strain suffered from pharyngitis and pyrexia without any muscle involvement (Shaw, Melnick & Curnen, 1950).

b. Bornholm Disease

This disease, first described as pleurodynia or Bamle disease until the classic description by Sylvest (1934) of an outbreak on the island of Bornholm, is also caused by C viruses, predominantly those of the B group. The clinical picture of this disease is well known; it is associated with severe attacks of muscle pain generally affecting the lower thorax and upper abdomen, fever, pharyngitis, and sometimes a dry pleurisy. Atypically, myalgia may be either absent or mild, or definite meningitis may occur. C viruses have often been isolated from this disease since they were first shown, by Weller, Enders, Buckingham & Finn (1950) in the USA, to be definitely involved. They have been found in outbreaks in Switzerland, Holland, France, Israel, Sweden, Australia and in Great Britain (Thélin & Wirth, 1951; Atkinson, Dineen & Robertson, 1951; and by others). The virus has also been isolated from muscle biopsy specimens in human cases and these have shown the typical myositis associated with virus infection in suckling mice (Lépine, Desse & Sautter, 1952).

In 1951 many outbreaks of the disease were reported in the United Kingdom (*British Medical Journal*, 1951; McNeish & Stewart, 1952). From virus isolations and serological investigations it was found that these outbreaks were associated with two serologically different strains of C virus. B3 strains were present in the north and Midlands of England and in Scotland (Disney, Howard, Wood & Findlay, 1953; Swain & Mitchell, 1953). This strain of Coxsackie virus had already been associated with Bornholm disease both in the USA (Lazarus, Johnston & Galbraith, 1952) and in Sweden (Gard & Johnsson, 1952). Another type of virus caused a large epidemic in Oxfordshire (Warin, Davies, Sanders & Vizoso, 1953). The Oxford strain was neutralized by antisera prepared against another strain associated with two cases of encephalitis and sporadic cases of Bornholm disease in the north-west of England but not by antisera against B strains 1–4 (Sanders, personal communication, 1952). That some cases of non-specific infection are caused by the same agent as the classical syndrome was illustrated by two outbreaks in the United Kingdom. In one of these the typical syndrome predominated, while in the other the younger children only had these symptoms, the older ones and the adults having a more non-specific illness without any severe muscle involvement (Brown, Liddle & Tobin, 1952; Bury & Tobin, 1952). Both these incidents were associated with the same type of C virus.

c. Aseptic Meningo-Encephalitis

Although in the outbreaks of Bornholm disease in the north-west of England, noted above, signs of meningitis were absent, the same virus was isolated from two cases of "aseptic meningitis" occurring independently of these outbreaks. In Sweden many of the cases due to the B3 type of virus had pleocytosis in the c.s.f., together with signs of meningeal involvement (Gabinus, Gard, Johnsson & Pöldre, 1952). Conversely Curnen, Shaw & Melnick (1949) showed that in an outbreak of cases of "aseptic meningitis" these and one other patient, who suffered from a typical attack of Bornholm disease, were infected with the B1 virus strain that Weller *et al.* (1950) isolated later from epidemic pleurodynia. The case of herpangina from which Stuart-Harris (personal communication, 1952) isolated a strain of Coxsackie virus also showed an increase in cells in the c.s.f., and Swedish workers have found the same thing in some of their cases (Johnsson

& Lindahl, 1953). Moreover, the A10 strain has been isolated from patients with meningo-encephalitis in Finland (Oker-Blom, 1952). Thus it seems that some cases of aseptic meningo-encephalitis, with or without the associated signs of herpangina or Bornholm disease, are caused by Coxsackie infection, and another example of this has been reported recently in Great Britain (Galpine & McCrae, 1953).

In the north-west of England, during the summer of 1951, four isolations of C virus were made from cases of this disease, in three of which these agents could have been the cause. These viruses were of three different serological types, two B3, one an unidentified B type and one an A2. The three strains of B viruses were associated with a small crop of cases of meningitis occurring in June and July, while the A group virus (A2) was isolated from a case of herpes simplex encephalitis in August. An A2 virus was also isolated from a case of meningitis in the Midlands of England in 1950 (Howard, personal communication, 1953). However, the vast majority of the cases occurring later in the north-west were not caused by any of these viruses. Cases occurring from August to October 1950 from which no Coxsackie virus was isolated were investigated for serological evidence of infection by these three strains and in only one case was a rise in titre demonstrated (Tobin, unpublished data, 1953). Presumably these cases of aseptic meningo-encephalitis were caused by either a mild type of poliomyelitis virus or by some as yet unidentified agent. An outbreak of aseptic meningo-encephalitis due to C virus, closely followed by one due to poliomyelitis virus, has been noted before (Curnen, 1952).

d. Poliomyelitis

Although Coxsackie viruses have been isolated from cases of poliomyelitis by Rhodes, Clark, Knowles, Shimada, Ritchie & Donohue (1950), Melnick, Kaplan, Zabin, Contreras & Larkum (1951), Stanley, Hayes & Dorman (1951), and by other workers, there is no definite evidence that they play much part in the symptomatology of this disease. The figures of Dalldorf (1952) indicated that, when the Coxsackie viruses were prevalent, paralytic disease was less than in those years when they were scarce. Certainly, during the widespread incidence of Coxsackie infection in England during the summer and autumn of 1951, the incidence of paralytic poliomyelitis was very much lower than in 1949–50 or 1952. But the Yale workers could find no relation between the incidence of paralysis and Coxsackie infection (Curnen, 1952).

Huebner *et al.* (1952) failed to isolate virus from 170 stool specimens from poliomyelitis cases in 1950 and, in 1951, discovered the virus in only 3 patients out of 76 from whom 246 specimens were obtained. All these isolations were from stools obtained after the patient had been in hospital for three to nine days, which suggested that hospitals might act as centres for the spread of infection of some C viruses. Gard & Johnsson (1952) could isolate Coxsackie virus only from 1% of 396 cases of poliomyelitis, although it was present in 3.7% of 411 cases of aseptic meningo-encephalitis.

However, in suckling mice interference has been demonstrated between the group B strains of C virus and poliomyelitis viruses (Dalldorf, 1951).

Only one strain, a type A4, has been isolated from a poliomyelitis case in Great Britain and that was found post mortem in the patient's blood (Howard, personal communication, 1953).

e. Other Diseases

Although C viruses have also been isolated from other conditions diagnosed as summer grippe, etc., their role in these conditions is unknown, although many of these syndromes could have been the atypical form of Bornholm disease. In Great Britain, Forrester & Tobin (1951) isolated an A1 type virus from a case with an atypical clinical syndrome but were unable to decide definitely what part it played in the patient's disease. A1 strains of C virus have not been associated with herpangina and their clinical role is in doubt. They also behave differently from the other A strains in laboratory animals (Beeman, Huebner & Cole, 1952) and should probably be placed in a group by themselves. A group B virus has also been found associated with a non-specific disease in the Midlands of England (Kenyon, McCrae, Dodd & Galpine, 1952).

Treatment of infections with C virus is purely symptomatic, as they are completely resistant to antibiotics and chemotherapy.

3. Epidemiology

The common source of the Coxsackie viruses is faeces, therefore it is not surprising that they have been isolated from sewage and flies (Melnick, Shaw & Curnen, 1949); in fact virus can survive both in flies (Melnick & Penner, 1952) and in cockroaches (Fischer & Syverton, 1951) for two weeks and these insects may act as vectors for the spread of disease. After infection, virus may be excreted by patients for as long as two to three months but usually it has disappeared from the faeces in from one to four weeks. The ready isolation of virus from sewage suggests that in epidemic periods many people may be excreting virus (Clark, Knowles, Shimada, Rhodes, Ritchie & Donohue, 1951). Viruses have also been found in throat washings and have been isolated definitely from the c.s.f. on one occasion (Gabinus *et al.* 1952). Most isolations have been made during the summer and autumn, a time when herpangina and Bornholm disease are likely to be prevalent. Many of the outbreaks of these diseases can be explained by person-to-person contact, but the more widespread distribution of some of the cases would indicate some other form of dissemination, suggesting that the epidemiology is similar to that of poliomyelitis. Although Findlay (1952) has suggested that droplet infection is the normal method of spread from case to case, it is more probable that these viruses are spread in the same manner as other enteric agents, i.e., by infected fomites and by food and water. Their incidence should therefore be reduced by a high standard of personal and communal hygiene. The incubation period of C virus infections is generally about 2–5 days but rarely may be as long as 9–14 days.

4. Diagnosis of Coxsackie Infections

Diagnosis of infections by these viruses depends on their isolation and the subsequent demonstration of an increase of antibody in the patient's serum against the infecting strain. As the virus may be excreted for some weeks in the faeces, mere isolation of virus does not necessarily indicate that the disease occurring at the same time is due to Coxsackie infection and, unless confirmed by serological means, it is impossible to be precise as to the time of infection. Serological diagnosis alone is not a practical routine proposition, owing to the multiplicity of virus types.

a. Virus Isolation

In epidemics of Coxsackie infection the virus can readily be isolated from the throat washings during the first few days of illness and from the faeces during the first week and often later. The usual procedure is to emulsify faeces in saline, spin down the coarse particles in a horizontal centrifuge, treat the supernatant fluid overnight with ether to reduce the bacterial count and then spin it in the angle centrifuge to deposit any surviving bacteria and fine debris. After treatment with streptomycin and penicillin, the extract is inoculated either subcutaneously, intraperitoneally or intracerebrally into suckling mice. Methods of concentrating virus by high speed centrifugation (Clark, Rhodes, Knowles, Ritchie & Donohue, 1950) or by precipitation with ammonium sulphate (Gard & Johnsson, 1952) have also been used. If virus is present, signs of illness appear within 3 to 6 days with the A group and within 5 to 12 days with the B group, depending on the titre and the strain involved. Passage can then be continued either by using brain or carcass in the case of group B, or by the use of carcass in the case of group A viruses. On the whole the infective titres obtained with the latter are higher, as some group B viruses may infect mice at titres of only 10^{-3} and even after many passages may not reach titres above 10^{-5}. Virus is identified first by the pathological lesions it produces and then by means of neutralization (Melnick & Ledinko, 1950), complement fixation (Kraft & Melnick, 1950; Casals, Olitsky & Murphy, 1949) or cross-protection tests (Melnick, Clarke & Kraft, 1950). The most satisfactory method is by neutralization, as some B viruses never rise to a high enough titre to give a satisfactory complement-fixing antigen. During epidemics the virus can be isolated from faeces in over 80% of cases infected by the A strains and in approximately 30–40% of B infections (Gard & Johnsson, 1952; and see Table II).

TABLE II. ISOLATION OF COXSACKIE VIRUSES IN THE NORTH-WEST OF ENGLAND
(August 1950 to March 1952)

Disease	Number of cases from which faeces were examined	Number of "C" viruses isolated	Percentage isolated	Types isolated	Date isolated
Sporadic cases of myalgia	5	0	0	—	—
Epidemic Bornholm disease	9	3	33.3	B3	June 1951 July 1951 July 1951
Aseptic meningo-encephalitis	47	4	8.5	B3 B3 B ? A2	June 1951 July 1951 June 1951 August 1951
Other diseases	26	1	3.8	A1	August 1950

b. Serological Diagnosis

In human infection with C virus a rise in neutralizing antibody can be demonstrated, provided that the first sample of serum is taken within the first few days, and the second ten days or more afterwards. Antibodies appear on the fifth to ninth day and increase rapidly, in many cases to their maximum within a few days, and persist for at least a year with the group B viruses and for longer periods with those of group A. The usual method for demonstrating antibodies is to test falling dilutions of serum against a constant amount of virus, usually 100 ID_{50}, as the neutralization index may be very high as soon as antibodies appear. The usual titres of antibody obtained with this dose of virus vary between 1 in 50 and 1 in 1,250, depending on the infecting strain of virus. Although complement fixation has been demonstrated between virus antigen and human antibody in cases of Coxsackie infection, this test is not suitable for diagnostic work, as often there may be a rise in heterologous antibody as well as in the homologous one without any corresponding rise in neutralizing antibody. Whether this is due to a group antigen contained in all Coxsackie viruses (Beeman & Huebner, 1952) or whether it is in the nature of an anamnestic reaction is not known (Kraft & Melnick, 1952). Complement-fixing antibodies do not persist for the same length of time as neutralizing ones and often disappear or drop to a low level within six months (Curnen, 1952).

5. Summary

The Coxsackie viruses are widespread in all countries in which search has been made for them and can cause definite illness, of which the main diseases are herpangina, Bornholm disease, and aseptic meningo-encephalitis. They are rarely, if ever, fatal but cause a great deal of incapacity and mild sickness, especially in children. Man appears to be their only host and their spread is similar to that of other enteric viruses. Serologically they form a heterogeneous group of at least 16 different types. At least six of these (A1, A2, A4, A10, B3 and another untyped B variety) have been found in the United Kingdom. However, many of the problems of the general ecology of this group of viruses remain to be elucidated. Luckily these viruses can be isolated and propagated in a readily available laboratory animal, that is, the suckling mouse.

ACKNOWLEDGEMENT

Thanks are due to Dr G. Dalldorf for typing some of the strains isolated.

REFERENCES

Atkinson, N., Dineen, J. & Robertson, J. S. (1951) *Aust. J. exp. Biol. med. Sci.* **29,** 463
Beeman, E. A. & Huebner, R. J. (1952) *J. Immunol.* **68,** 663
Beeman, E. A., Huebner, R. J. & Cole, R. M. (1952) *Amer. J. Hyg.* **55,** 83
Briefs, A., Breese, S. S., Warren, J. & Huebner, R. J. (1952) *J. Bact.* **64,** 237
British Medical Journal (1951) Correspondence in: *Brit. med. J.* **2,** 1089
Brown, C. M., Liddle, D. C. & Tobin, J. O'H. (1952) *Lancet*, **1,** 445
Bury, H. S. & Tobin, J. O'H. (1952) *Lancet*, **2,** 267
Casals, J., Olitsky, P. K. & Murphy, L. C. (1949) *Proc. Soc. exp. Biol., N.Y.* **72,** 636
Clark, E. M., Knowles, D. S., Shimada, F. T., Rhodes, A. J., Ritchie, R. C .& Donohue, W. L. (1951) *Canad. J. publ. Hlth*, **42,** 103
Clark, E. M., Rhodes, A. J., Knowles, D. S., Ritchie, R. C. & Donohue, W. L. (1950) *Canad. J. publ. Hlth*, **41,** 51
Curnen, E. C. (1952) In : International Poliomyelitis Congress. *Poliomyelitis : papers and discussions presented at the Second International Poliomyelitis Conference*, p. 121. Lippincott, London

Curnen, E. C., Shaw, F. W. & Melnick, J. L. (1949) *J. Amer. med. Ass.* **141,** 894
Dalldorf, G. (1950) *Bull. N. Y. Acad. Med.* **26,** 329
Dalldorf, G. (1951) *J. exp. Med.* **94,** 65
Dalldorf, G. (1952) In: International Poliomyelitis Congress. *Poliomyelitis: papers and discussions presented at the Second International Poliomyelitis Conference,* p. 111. Lippincott, London
Dalldorf, G. & Gifford, R. (1952) *J. exp. Med.* **96,** 491
Dalldorf, G. & Sickles, G. M. (1948) *Science,* **108,** 61
Disney, M. E., Howard, E. M., Wood, B. S. B. & Findlay, G. M. (1953) *Brit. med. J.* **1,** 1351
Findlay, G. M. (1952) Discussion in: International Poliomyelitis Congress. *Poliomyelitis: papers and discussions presented at the Second International Poliomyelitis Conference,* p. 135. Lippincott, London
Findlay, G. M. & Howard, E. M. (1950) *Brit. med. J.* **1,** 1233
Fischer, R. G. & Syverton, J. T. (1951) *Amer. J. trop. Med.* **31,** 238
Forrester, R. M. & Tobin, J. O'H. (1951) *Lancet,* **2,** 663
Gabinus, O., Gard, S., Johnsson, T. & Pöldre, A. (1952) *Arch. ges. Virusforsch.* **5,** 1
Galpine, J. F. & McCrae, A. D. (1953) *Lancet,* **1,** 372
Gard, S. & Johnsson, T. (1952) *Acta path. microbiol. scand.* Suppl. No. 93, p. 332
Gifford, R. & Dalldorf, G. (1951) *Amer. J. Path.* **27,** 1047
Himmelweit, F., Findlay, G. M. & Howard, E. M. (1950) *Brit. J. exp. Path.* **31,** 809
Huebner, R. J., Beeman, E. A., Cole, R. M., Bergelman, P. M. & Bell, J. A. (1952) *New Engl. J. Med.* **247,** 285
Johnsson, T. & Lindahl, J. (1953) *Arch. ges. Virusforsch.* **5,** 96
Kenyon, H., McCrae, A. D., Dodd, R. J. & Galpine, J. F. (1952) *Lancet,* **2,** 153
Kraft, L. M. & Melnick, J. L. (1950) *J. exp. Med.* **92,** 483
Kraft, L. M. & Melnick, J. L. (1952) *J. Immunol.* **68,** 297
Kravis, L. P., Hummeler, K. & Sigel, M. M. (1952) *Bact. Proc.,* p. 105 [Abstract]
Lazarus, A. S., Johnston, E. A. & Galbraith, J. E. (1952) *Amer. J. publ. Hlth,* **42,** 20
Lépine, P., Chaumont, L. & Blusson, J. (1952) *Ann. Inst. Pasteur,* **82,** 648
Lépine, P., Desse, G. & Sautter, V. (1952) *Bull. Acad. nat. Méd., Paris,* **136,** 66
McNeish, W. W. W. & Stewart, C. (1952) *Brit. med. J.* **1,** 744
Melnick, J. L. (1951) *Annu. Rev. Microbiol.* **5,** 309
Melnick, J. L., Clarke, N. A. & Kraft, L. M. (1950) *J. exp. Med.* **92,** 499
Melnick, J. L., Kaplan, A. S., Zabin, E., Contreras, G. & Larkum, N. G. (1951) *J. exp. Med.* **94,** 471
Melnick, J. L. & Ledinko, N. (1950) *J. exp. Med.* **92,** 463
Melnick, J. L. & Penner, L. R. (1952) *J. exp. Med.* **96,** 255
Melnick, J. L., Shaw, E. W. & Curnen, E. C. (1949) *Proc. Soc. exp. Biol., N.Y.* **71,** 344
Oker-Blom, N. (1952) Discussion in: *Acta path. microbiol. scand.* Suppl. No. 93, p. 338
Pappenheimer, A. M., Kunz, L. J. & Richardson, S. (1951) *J. exp. Med.* **94,** 45
Parrott, R. H., Ross, S., Burke, F. G. & Rice, E. C. (1951) *New Engl. J. Med.* **245,** 275
Peers, J. H., Ransom, S. E. & Huebner, R. J. (1952) *J. exp. Med.* **96,** 17
Rhodes, A. J., Clark, E. M., Knowles, D. S., Shimada, F. T., Ritchie, R. C. & Donohue, W. L. (1950) *Canad. J. publ. Hlth,* **41,** 183
Shaw, E. W., Melnick, J. L. & Curnen, E. C. (1950) *Ann. intern. Med.* **33,** 32
Slater, E. A. & Syverton, J. T. (1950) *Proc. Soc. exp. Biol., N.Y.* **74,** 509
Stanley, N. F., Dorman, D. C. & Ponsford, J. (1953) *Aust. J. exp. Biol. med. Sci.* **31,** 31
Stanley, N. F., Hayes, L. & Dorman, D. C. (1951) *Aust. J. exp. Biol. med. Sci.* **29,** 367
Sulkin, S. E., Wallis, H. C. & Donaldson, P. (1952) *J. infect. Dis.* **91,** 290
Swain, R. H. A. & Mitchell, R. G. (1953) *Brit. med. J.* **1,** 1354
Sylvest, E. (1934) *Epidemic myalgia: Bornholm disease.* Milford, London
Thélin, F. & Wirth, J. (1951) *Rev. méd. Suisse rom.* **71,** 44
Warin, J. F., Davies, J. B. M., Sanders, F. K. & Vizoso, A. D. (1953) *Brit. med. J.* **1,** 1345
Weller, T. H., Enders, J. F., Buckingham, M. & Finn, J. J. (1950) *J. Immunol.* **65,** 337
Zahorsky, J. (1920) *Sth. med. J., Nashville,* **13,** 871

* * *

THE COMMON COLD

C. H. ANDREWES M.D. F.R.C.P. F.R.S.

National Institute for Medical Research
Mill Hill, London

1 The Salisbury experiments
2 Properties of the virus
3 Natural transmission of infection
4 Prevention and cure of colds
5 Prospects for prevention of colds
References

The common cold is a ubiquitous but, from the laboratory point of view, an elusive virus disease. It probably represents a definite entity but similar symptoms may be produced by other infections with longer incubation periods—febrile catarrh (Stuart-Harris, Andrewes & Smith, 1938) or acute undifferentiated respiratory disease (A.R.D.) and atypical pneumonia (Dingle, 1947). The common cold was shown by Kruse (1914) and by Foster (1917) to be transmissible by means of bacteria-free filtrates and to be, almost certainly, caused by a virus. Progress in its study has been hampered by the fact that man and the chimpanzee (cf. Dochez, 1933) are the only known susceptible animals. The infection leaves behind it little or no lasting immunity, so that most people, in temperate climates at least, suffer two or more attacks every year, usually in the winter months (cf. review by Greenspan, 1943). Members of isolated communities may long remain free from colds but, when they ultimately do make contact with "civilization", they are fairly sure to develop colds of particular severity. The best-documented accounts are those of Paul & Freese (1933), who made observations on Spitzbergen, and those of Heinbecker & Irvine-Jones (1928) amongst Polar Eskimos. It is likely therefore that such immunity as most people have is largely determined by repeated contact with small doses of virus.

Attempts to implicate bacteria in the primary causation have not been successful, though all agree that these play a part, as secondary invaders, in making the later stages more unpleasant, particularly when sinusitis follows a simple cold.

1. The Salisbury Experiments

In 1946 a team of workers under the Medical Research Council, with the collaboration of the Ministry of Health, initiated a fresh attack on the problem at the Harvard Hospital, Salisbury, in southern England. This was primarily an attempt to apply newer virus techniques to the discovery of a method of recognizing and titrating the virus in the laboratory. The plan involved, as a means to this end, the use of many human volunteers and over 3,000 were so used. It was particularly necessary to plan the experiments in such a way that neither volunteers nor clinical observers knew whether the inocula tested on them were cold filtrates or harmless "control" fluids. It was also necessary that the volunteers should be isolated from outside contacts for the ten days' duration of each experiment. The techniques successfully applied in this study have been described in detail (Andrewes, 1949, 1950).

In the first six and a half years of this experiment, the main objective was not attained. It was not found possible to infect with colds any of 20 animal species, including several species of primate; nor even to demonstrate survival of virus in their nasal passages. The virus was not cultivated in fertile eggs, either in the yolk sac, amniotic or allantoic cavities, nor was it possible to confirm any of several claims to have so cultivated the virus (Kneeland, Mills & Dochez, 1936; Topping & Atlas, 1947; Pollard & Caplowitz, 1947; Ward & Proctor, 1950). The virus survived in a number of instances in tissue cultures of human embryonic nasal epithelium, but multiplication and serial propagation were not proved.

Recent experiments indicate that the common cold virus can be serially propagated in cultures of human embryonic lung (Andrewes, Chaproniere, Gompels, Pereira & Roden (1953)).

2. Properties of the Virus

A number of facts were established concerning the properties of the common cold virus. It was readily preserved—for two years or more—at −76° C. in dry ice, but it did not survive attempts at freeze-drying. Filtration through gradocol membranes indicated that it was certainly less than 70 mμ. in diameter and possibly as small as 30 mμ. It was unaffected by penicillin and other antibiotics, but was readily inactivated in vitro by 20% ethyl ether. Colds were produced by 1 in 100 dilutions of nasal washings, sometimes by 1 in 1,000, never by 1 in 10,000.

The most active nasal washings, filtered or not, would produce colds in 55% of women and in 43% of men. Age, within the age-group 18 to 40, did not affect these percentages, nor apparently did the period which had elapsed since the subject's last cold nor the time of year of the experiments. These conclusions are tentative ones, based on the first few years of the study. The incubation period was usually two to three days from the time of intranasal instillation of virus. Initial symptoms consisted, in two-thirds of the cases, of soreness at the back of the throat. A cold was not diagnosed in the absence of increased nasal discharge. About a third of the subjects had a little cough, sooner or later. These experimental colds were mostly mild and cleared up within a week.

Virus was plentifully present in nasal secretions of one volunteer 24 hours, but not 12 hours, after infection: symptoms began in this instance only after 48 hours. This and other evidence suggests that a subject may well be infectious before symptoms appear. An experience with chimpanzees (Long, Bliss & Carpenter, 1932) supports this idea. Addition of various adjuvants to inocula failed to increase their efficacy. Chilling by various means, including the wearing of wet socks, could not be shown to produce colds by itself nor to lower resistance to a small dose of virus.

3. Natural Transmission of Infection

Early attempts at Harvard Hospital to produce natural contact infection gave a very low incidence of "takes"; this made the study of the most effective route of transfer very difficult. Knowledge that members of isolated communities

become highly susceptible to colds suggested that it would be worth trying experiments on such persons. Twelve students were accordingly isolated on Eilean nan Ròn, off the north coast of Scotland, for three months (Andrewes, Lovelock & Sommerville, 1951). People artificially infected with common colds were then brought on to the island and exposed under various conditions to the supposedly hypersusceptible students. Unexpectedly none of the latter developed colds. This failure may have been due to a change induced in the common cold virus by artificial nose-to-nose transfer in the months preceding this experiment, so that, though the virus would still produce colds, it would not pass naturally by contact. This suggestion is supported by the observation that a natural cold in a crofter who was subsequently brought on to the island did spontaneously "jump" to three of the students.

Following this experiment, renewed contact experiments were carried out at Salisbury, using children with natural or "wild" colds as sources of infection. There is evidence (Lidwell & Sommerville, 1951) that colds in children spread more readily to others than do colds in adults. Experiments showed that in two hours of normal social intercourse about 10% of adults would pick up infection from such children. Exposure to inhalation of droplet nuclei (cf. Wells & Wells, 1936) would also produce some infections; but contact with contaminated environment gave less conclusive results.

Previous work had indicated that infected nasal secretions might very readily be widely spread around manually. It is another matter whether virus survives such transfer and can subsequently infect others. Several tests indicated that cold virus does not readily survive drying (Lovelock, Porterfield, Roden, Sommerville & Andrewes, 1952). Persons used handkerchiefs contaminated with cold secretions and then briefly dried; few of them developed colds. None did so when infected secretions were painted on the outside of the nares. Infected gauze packed into the nares would produce colds, but much less effectively if dried first. Spraying infected fluid was a less effective way of transferring a cold than administering drops: virus in fine droplets is of course subject to rapid drying.

Epidemiological studies in a valley near Salisbury (Lidwell & Sommerville, 1951) indicated that the chances that an individual would catch a cold when infection was introduced into a household were only 1 in 5. Even so, the risk is greater than that in the experiments just described where exposure was for two hours, but that is not inconsistent, seeing that natural exposure in the valley families was for a matter of days. Adults in houses where there was a school-age child caught twice as many colds as others who lacked this regular house-contact with children; females had more colds than did males—a result in conformity with the greater susceptibility of women to experimental infection, already mentioned.

4. Prevention and Cure of Colds

Bacterial vaccines have been widely used for prophylaxis of colds but the most carefully controlled trials (Ferguson, Davey & Topley, 1927; Diehl, Baker & Cowan, 1938) have failed to produce evidence in their support. Within recent years antihistaminic drugs have been claimed as valuable agents for aborting or curing colds; but numerous papers have since appeared contradicting these claims. In a Report of the Medical Research Council (1950) on the subject is included reference to experiments at Salisbury, where two antihistaminics proved useless in preventing colds amongst volunteers, even when administered before infection and continued thereafter. Modern advice as to local medication of colds is to leave the nose well alone: it is much easier to do harm than good.

5. Prospects for Prevention of Colds

Measures of air-sterilization or air-hygiene, including the use of ultra-violet irradiation and disinfectant mists, have not yet shown much promise as methods of preventing colds. Indeed it is possible that, if they were successful in their immediate aim, these techniques might be actually harmful: people not regularly exposed to small doses of virus might become abnormally susceptible so that when they ultimately did catch colds they would suffer from a severe instead of a relatively benign infection. Vaccines of orthodox types also seem unpromising in view of facts recorded above and of the brief duration of natural immunity.

Prospects are, however, by no means hopeless. People apparently vary in their susceptibility to colds from time to time. Only about 50% of volunteers at Salisbury catch colds when inoculated with potent material; yet these same volunteers are liable to catch colds at other times. Presumably there is variation in some as yet elusive factor determining resistance; knowledge of its nature might give a very useful clue. Again, there are numerous instances of attenuation of viruses by propagation in an unaccustomed host or culture; success in growing the virus elsewhere than in the human nose might bring it to a state in which it could be used as a locally applied vaccine.

REFERENCES

Andrewes, C. H. (1949) *Lancet*, **1**, 71
Andrewes, C. H. (1950) *New Engl. J. Med.* **242**, 235
Andrewes, C. H., Chaproniere, D. M., Gompels, A. E. H., Pereira, H. G. & Roden, A. T. (1953) *Lancet*, **2**, 546
Andrewes, C. H., Lovelock, J. E. & Sommerville, T. (1951) *Lancet*, **1**, 25
Diehl, H. S., Baker, A. B. & Cowan, D. W. (1938) *J. Amer. med. Ass.* **111**, 1169
Dingle, J. H. (1947) *Trans. Coll. Phycns Philad.* **15**, 113
Dochez, A. R. (1933) *Medicine, Baltimore*, **12**, 245
Ferguson, F. R., Davey, A. F. C. & Topley, W. W. C. (1927) *J. Hyg., Camb.* **26**, 98
Foster, G. B. (1917) *J. infect. Dis.* **21**, 451
Greenspan, F. (1943) *Arch. Pediat.* **60**, 90; 149; 208
Heinbecker, P. & Irvine-Jones, E. I. M. (1928) *J. Immunol.* **15**, 395
Kneeland, Y., jr, Mills, K. C. & Dochez, A. R. (1936) *Proc. Soc. exp. Biol., N.Y.* **35**, 213
Kruse, W. (1914) *Münch. med. Wschr.* **61**, 1547
Lidwell, O. M. & Sommerville, T. (1951) *J. Hyg., Camb.* **49**, 365
Long, P. H., Bliss, E. & Carpenter. H. M. (1932) *Johns Hopk. Hosp. Bull.* **51**, 278
Lovelock, J. E., Porterfield, J. S., Roden, A. T., Sommerville, T. & Andrewes, C. H. (1952) *Lancet*, **1**, 657
Medical Research Council: Special Committee (1950) *Brit. med. J.* **2**, 425
Paul, J. H. & Freese, H. L. (1933) *Amer. J. Hyg.* **17**, 517
Pollard, M. & Caplowitz, C. D. (1947) *Science*, **106**, 243
Stuart-Harris, C. H., Andrewes, C. H. & Smith, W. (1938) *Spec. Rep. Ser. med. Res. Coun., Lond.* No. 228
Topping, N. H. & Atlas, L. T. (1947) *Science*, **106**, 636
Ward, T. G. & Proctor, D. F. (1950) *Amer. J. Hyg.* **52**, 91
Wells, W. F. & Wells, M. W. (1936) *J. Amer. med. Ass.* **107**, 1698; 1805

THE VIRUSES OF EPIDEMIC INFLUENZA

A. ISAACS M.B. Ch.B.

World Influenza Centre
National Institute for Medical Research
Mill Hill, London

1 Introduction
2 Antigenic characters of influenza viruses
3 Epidemiology of influenza in relation to antigenic characters of the viruses
4 Prevention of influenza
References

1. Introduction

In 1933, Smith, Andrewes and Laidlaw, working at the National Institute for Medical Research, Hampstead, London, isolated influenza virus by inoculating human garglings into ferrets. So many lectures and articles on influenza have commenced with words like these that it is easy to pay lip service to the earlier fundamental experiments on influenza and then to forget them in the welter of exciting publications with which we are flooded today. During the last few years knowledge of the influenza viruses and their behaviour under carefully controlled laboratory conditions has increased enormously and today they are the organisms most used as a model for studying the multiplication of animal viruses. Yet we must go back to the early years after 1933 to look for the important advances in our knowledge of influenza the disease.

In their first paper on influenza, Smith, Andrewes & Laidlaw (1933) described how they produced an influenza-like disease in ferrets by intranasal inoculation of filtrates from cases of human influenza. The first strain isolated, and known today as the WS strain, was maintained by passage of suspensions of ferret turbinates. The disease was shown to be contagious in experimental ferrets which had therefore to be strictly isolated. On recovery from infection, ferrets were immune to challenge, and their sera and human convalescent sera neutralized the virus. Considerable cross-immunity was demonstrated between the WS strain and a strain of the virus of swine influenza (Shope, 1931). Transmission of influenza virus to mice (Andrewes, Laidlaw & Smith, 1934; Francis, 1934) made the study of influenza more practicable, since sera could now be tested for antibody by a simple neutralization test. Growth of virus was next obtained in tissue culture (Smith, 1935; Francis & Magill, 1935) and on the chorio-allantoic membrane of the developing chick embryo (Smith, 1935). Meanwhile, Laidlaw (1935) was intrigued with the relationship between human and swine influenza viruses. He postulated that the influenza pandemic of 1918–19 was due to a virus of the swine influenza type which at that time became adapted to the pig, in which it has since persisted unchanged, while undergoing progressive modifications during successive epidemics in human beings. This fascinating hypothesis was strongly supported by Shope (1936) amongst others.

Complement-fixation tests were developed for the serological diagnosis of influenza (Smith, 1936; Fairbrother & Hoyle, 1937), and the latter authors found (Hoyle & Fairbrother, 1937) that the antigen responsible for complement fixation in their tests could be separated from the virus elementary bodies by centrifugation. The elementary bodies had been shown by the filtration studies of Elford, Andrewes & Tang (1936) to have a diameter of 80–120 mμ. and the smaller complement-fixing (soluble) antigen was thought by Hoyle & Fairbrother (1937) to be a product of virus multiplication. In support of this they found that convalescent ferret sera contained neutralizing and complement-fixing antibody, whereas after vaccination ferret sera contained only neutralizing antibody. An important contribution to our knowledge of influenza the disease was the correlation (Stuart-Harris, Andrewes & Smith, 1938) between the clinical and laboratory findings in cases of influenza and the delineation of febrile catarrh as a distinct and unrelated entity. The independent isolation of strains of influenza virus serologically distinct from earlier strains, and afterwards called influenza B, by Francis (1940) and Magill (1940) marked an important step in epidemiological knowledge.

Influenza research now took quite a different turn and the ferret and mouse were replaced by the egg and test-tube. Burnet (1940) showed that influenza virus could be readily cultivated by amniotic inoculation of chick embryos and its growth recognized by the occurrence of large vacuolated epithelial cells in tracheal smears. There followed the fundamental observations of Hirst (1941) and of McClelland & Hare (1941) that influenza virus agglutinated red blood cells of a number of animal species, a finding which has been of the greatest theoretical and practical importance. Today a new strain of influenza virus is usually isolated by amniotic inoculation of chick embryos, then adapted to the allantoic cavity (Nigg, Crowley & Wilson, 1940) and studied thereafter mainly by agglutination techniques. The great simplicity of the newer techniques has led to many important theoretical advances in our knowledge of influenza viruses.

In the past ten years many different aspects of influenza viruses have been investigated. The fundamental researches of Hirst (1942) on the agglutination and the enzymic-like elution from red cells stimulated Burnet and his colleagues in Australia to study this reaction as a model of the processes of invasion of susceptible cells by influenza viruses (Burnet, 1951). The influenza viruses show many examples of virus variation, and such phenomena as the development of neurotropism (Stuart-Harris, 1939), the O–D variation (Burnet & Bull, 1943) and the adaptation of virus to kill mice following intranasal inoculation (see, e.g., Hirst, 1947; Wang, 1948; Davenport & Francis, 1951; Chu, 1951) have been fruitful fields of study. The Henles and their colleagues in Philadelphia have made elegant studies of the interference phenomena associated with influenza viruses (Henle, 1950) in an attempt to analyse the processes of virus multiplication, while in Australia a different approach to the same problem comes from the fascinating work on recombination (Burnet, Fraser & Lind, 1953) in which it was found that, following simultaneous infection with two influenza strains, the products of infection showed a cross-over of some of their inheritable characters. Fazekas de St Groth and co-workers in Australia have made detailed investigations of the process of experimental influenza in mice (Fazekas de St Groth, 1950), and in Great Britain the technique (Dawson &

Elford, 1949) of adsorbing influenza virus on red-cell ghosts for electron microscopy has been most useful and has led to the striking demonstration of filamentous forms of influenza virus (Chu, Dawson & Elford, 1949). In the space available it would obviously be impossible to consider more than a single avenue of recent research on influenza viruses and it was decided to select that avenue which has been the centre of interest in Great Britain, that is, the epidemiology of influenza, particularly in relation to the antigenic characters of the influenza viruses. Many advances have occurred in this field since the setting up by the World Health Organization of an international organization for the study of influenza. This arose from the realization that the global epidemiology of influenza and the effectiveness of any preventive measures required study on a world-wide basis, a consideration which has produced important changes in our outlook on this disease.

2. Antigenic Characters of Influenza Viruses

There are two main serological types of influenza viruses, A and B, and a third type has been recently proposed. The influenza A viruses are by far the most important cause of major epidemics of influenza. Serological heterogeneity among them was early recognized (Magill & Francis, 1936, 1938; Andrewes, 1937; Smith & Andrewes, 1938; Burnet & Clark, 1942), although some of the differences found between strains may have arisen as a result of animal passage. Since eggs exclusively have been used for virus culture the conception has arisen (Hirst, 1943) that strains isolated from one epidemic are serologically fairly uniform, while those isolated from different epidemics are more readily distinguishable. As a result it seems that the influenza A viruses have been undergoing a progressive series of antigenic changes (Andrewes, 1949; van der Veen & Mulder, 1950; Hilleman & Horsfall, 1952; Magill & Jotz, 1952), with continual emergence of new varieties. There are no facts, although there are many strongly held opinions, on whether the number of possible patterns is small, with antigenic changes pursuing a circular course, or infinite, with no limits set to the number of possible changes.

All influenza A viruses share a common soluble antigen which confers the serological characteristics on the group. Among the A viruses, if the swine viruses are excluded from consideration, three major sub-types can be recognized: the WS sub-type which was found up to 1933, the PR8 sub-type found from 1934 to 1946 and the so-called A prime (A′) sub-type found since 1946. Numerous subdivisions can be made within the PR8 and A′ sub-types and Hirst (1952) has used virus-absorbed sera in an attempt to clarify the problem. Among influenza A viruses isolated between 1933 and 1947 he has recognized at least six type-specific antigenic components which gradually replaced one another as the dominant antigenic component of strains isolated during these years.

In addition to the type-specific antigens the strains share a group-specific component, whose relationship to the soluble antigen is unknown, and quantitative variations in these antigens are thought to account for the numerous serological varieties of virus. Among the A′ strains, which have been more extensively investigated than earlier strains, the same picture holds. Strains isolated during one epidemic have tended to be serologically uniform (Chu, Andrewes & Gledhill, 1950), while there have been recognizable differences between A′ viruses isolated in different epidemic years (Hilleman, 1952; Isaacs, Gledhill & Andrewes, 1952).

The situation among influenza B viruses has been quite different. Although reports have appeared of numerous antigenic sub-types among B viruses the suggestion that these exist may arise largely from the techniques used. Workers who have used red cells of uniform sensitivity and sera from which normal inhibitors of agglutination have been removed find that strains recovered and passaged in eggs over the period 1943 to 1952 have shown a remarkable serological uniformity (Bozzo, 1952; Brans, 1952; Hennessen, 1952). Strains isolated in ferrets between 1940 and 1943, e.g., the Lee and TM strains, are serologically different from strains isolated over the period 1943 to 1952 and from one another. These differences may have resulted from the animal passages of the former strains, but even if they did not, it seems clear that influenza B strains recovered during the last ten years have shown far more serological uniformity than the A viruses.

A strain of virus, 1233, was isolated by Taylor (1949), and a second strain, serologically identical with 1233, was recovered by Francis, Quilligan & Minuse (1950). The latter authors proposed the name influenza C for this virus. In its laboratory characteristics it is clearly related to the influenza A and B viruses but its relationship to epidemic influenza has not yet been clearly established.

3. Epidemiology of Influenza in Relation to Antigenic Characters of the Viruses

The variability in severity of influenza epidemics has led some authors to suggest that there may be a virulence factor in the virus responsible for these differences. No such virulence factor has yet been identified and it has seemed profitable, therefore, to attempt to defend the alternative and simpler hypothesis that the antigenic characters of the influenza virus are sufficient to account for the variable epidemiology of influenza.

The first and most striking difference is that between epidemics of influenza A and B. It has formerly been suggested that influenza B epidemics tend to occur at four-to-six-year intervals, but recent experience suggests that influenza B may cause small scattered outbreaks each winter with occasionally more severe epidemics. During years of large epidemics of influenza A, the influenza B viruses tend to be overshadowed; thus, during the 1948–49 European epidemic of influenza A, 10 out of 90 strains received at the World Influenza Centre were influenza B (Chu *et al.* 1950) and during the 1950–51 epidemic 15 of 118 strains were B (Isaacs *et al.* 1952); the situation in 1953 appears to be similar. In the 1949–50 and 1951–52 winters, on the other hand, almost all the strains received were influenza B. In addition, many influenza B outbreaks have occurred simultaneously in widely scattered places where there were minimal opportunities for spread from one focus to another. Influenza B viruses may therefore be thought of as being constantly with us, of lower epidemic virulence than A viruses and responsible for localized outbreaks each winter. The relatively uniform antigenic behaviour of influenza B viruses isolated during the past ten years is in keeping with the behaviour of a virus which, like measles, can maintain itself in a human population by constantly finding a small number of susceptible hosts.

The situation with influenza A is quite different. In Europe, large-scale epidemics occur every two to three years. During a large epidemic it is not unusual to find a case-incidence of 5–10% of the population in a few weeks; many more people are probably infected subclinically (Hoyle, 1944). It is common to find large numbers of deaths reported as due to influenza and also, coinciding with the peak of the influenza epidemic, many more deaths in old people (Bradley, 1951). No other disease today affects such a high proportion of the population in such a short space of time, or causes such a sharp rise in the death-rate. After an epidemic of influenza A the virus, at least in an overt form, apparently disappears from the community. Between epidemics it may circle the world, finding susceptible individuals in different countries, although this theory does not satisfactorily account for the fact that outbreaks of influenza A in Europe, unlike B, do not occur every year; alternatively it may exist in recovered individuals in a latent form (cf. herpes simplex)[1] waiting to be activated (the basic virus of Andrewes, 1942). In a disease which affects such a high proportion of the population, any antigenic virus variants, however rare their appearance, would have a great survival advantage which would allow them to spread rapidly and replace earlier viruses. This does in fact represent the processes of infection with influenza A viruses. Thus the first A′ virus to be recognized was the Australian CAM (1946) strain, and within six months A′ viruses were isolated from almost every part of the world and seem to have replaced the earlier A viruses. Similarly, within six months of its first appearance in Australia and South Africa, viruses of the Liverpool sub-type had spread throughout the world.

A great deal of evidence has been accumulating to suggest that in addition to antigenic variation the influenza A viruses may have evolved a second mechanism which favours their survival. Van der Veen & Mulder (1950) described the recovery of influenza viruses with low avidity for antibody (Q viruses), in contrast to P viruses which showed high avidity for antibody. These observations have been confirmed in this laboratory and we have suggested that the influenza viruses may undergo a biphasic variation. In support of this it was possible to induce a change of P to Q virus in the laboratory by passage in immune serum, while the reverse change occurred after one or two passages in mice (Isaacs *et al.* 1952). There is some evidence that similar changes are occurring in nature. From a naturally occurring Liverpool sub-type virus in the P phase it was possible by growing it in immune serum to isolate Q variants with distinctive properties. One year later very similar variants were encountered in strains isolated from the Persian Gulf, and both P and Q variants were present side by side in the same " strains " (Isaacs, 1953). The readiness with which this change can be induced in the laboratory, and the frequency with which Q viruses are recovered naturally, suggest that when an influenza virus encounters a partially immune population a change to the Q phase may frequently result, possibly by some form of re-arrangement of antigenic components already present. This may be a commoner form of variation than that which results in the appearance of a new antigen—that is, antigenic variation.

Strong evidence for the importance of the antigenic characters of the virus in influencing the outcome of an influenzal outbreak comes from studies of the effects of a single virus on populations with different basic immunity to influenza virus (Pettit, Mudd & Pepper, 1936; Isaacs, Edney, Donnelley & Ingram, 1950). In the latter epidemic, in Ocean Island, a single strain of A′ virus attacked three different racial groups at once, caused severe influenza in 100% of native Gilbert and Ellice Islanders with no previous experience of the disease, moderate influenza in 10% of Europeans, but no disease in Chinese who presumably brought the virus with them to the Island. It is possible that the virus acquired " virulence " in its spread through highly susceptible individuals but the facts can be explained just as satisfactorily, without postulating any unknown factors, on the basis of the different past experiences of influenza virus in these populations.

How does epidemic influenza spread? The evidence that within a short time of its first appearance a new antigenic variant of influenza A can be recovered from all over the world leaves no doubt that there is true spread of influenza virus from one country to another. This does not, however, appear to be the only method of spread. There are a number of instances in which small influenza outbreaks have occurred in summer, died down, and then been followed by an outbreak in the same area in the following winter. In two instances, it has been shown that virus strains from the summer and winter epidemics were antigenically identical (USA Army Commission on Influenza, 1944; Isaacs & Andrewes, 1951). In addition there is evidence that in the winter outbreaks large numbers of people in widely scattered areas become infected almost simultaneously. These facts suggest that virus may enter an area at a time when conditions for epidemic spread are unfavourable and that it may become widely " seeded " in the population in a latent form which is " activated " in the following winter. This hypothesis is reminiscent of the epidemiology of swine influenza (Shope, 1943) and there is only circumstantial evidence to support it at present. Nevertheless, some of the epidemiological facts appear to require a hypothesis of this nature to explain them.

4. Prevention of Influenza

Prevention of influenza is frequently equated with vaccination. Early in the investigation of influenza it was shown that vaccination would protect ferrets (Smith, Andrewes & Laidlaw, 1935) and mice (Francis & Magill, 1935) against infection. The first trials of vaccine in man were carried out soon afterwards (Francis & Magill, 1936) and in this and subsequent trials it was shown unequivocally that following vaccination a rise in serum antibody to the virus occurred. The correlation with protection has been less impressive. One reason for this, as Francis (1941) pointed out, is that circulating antibody does not control immunity but that antibody in nasal secretions is probably the important factor in protection against influenza. Following vaccination, there is a rise in the antibody titre in nasal secretions (Francis, Pearson, Sullivan & Brown, 1943; Mulder, Brans & Hers, 1952), and the careful studies of Fazekas de St Groth and co-workers on experimental influenza in mice (Fazekas de St Groth, 1950) underline this and suggest methods by which nasal antibody might be increased in amount (Fazekas de St Groth, Donnelley & Graham, 1951).

In spite of its uneconomical nature, antigenically speaking, parenteral injection with influenza virus vaccines has had its successes. In a large trial carried out in the USA in 1943

[1] See Burnet, p. 173 of this Bulletin.

and involving approximately 6,000 vaccinated and control individuals, the incidence of influenza was reduced from 7.1% to 2.2% by vaccination (USA Army Commission on Influenza, 1944). On the other hand, in similar trials in 1947 in the USA and in Great Britain, vaccination was ineffectual, and this was shown to be due to the fact that the vaccines given contained strains only very distantly related antigenically to the A′ virus which caused the epidemic. The appearance of new antigenic variants anywhere in the world thus requires the earliest possible recognition in order that new strains can be rapidly incorporated into current vaccines. The use of adjuvants to increase and sustain the antigenic response to vaccination (Salk, Laurent & Bailey, 1951) should be of value provided that the risks attending the use of adjuvants can be completely eliminated, as has been suggested (Salk, Bailey & Laurent, 1952). Recently-recovered strains of influenza virus have been described as poorer antigens than those isolated earlier and this may be due to their having been isolated and maintained in eggs. Isaacs & Sampaio (1953) have shown that adaptation to mice greatly increases the antigenicity of a virus for mice; apparently it also increases its antigenicity for human beings (Isaacs & Sampaio, work to be published) and it may be worth while therefore considering adapting newly isolated strains to mice before using them in vaccines.

In spite of the large amount of investigation which the influenza viruses have undergone in the laboratory and in the field, many important epidemiological questions remain unanswered. They are a challenge to research workers everywhere, and the known versatility and variability of the virus should be sufficient guarantee that the problems cannot be lightly neglected. In the preventive field there is little room for satisfaction, although the steady progress made to date should encourage further investigation. It remains to be seen whether more progress will eventually come from the frontal attack of the epidemiologist or the lateral attacks of those interested in the more academic aspects of influenza virus growth. Perhaps the virus will finally be outwitted only when the imagination of man has discovered the innermost secrets of its way of life.

References

Andrewes, C. H. (1937) *Brit. med. J.* **2,** 513
Andrewes, C. H. (1942) *Proc. R. Soc. Med.* **36,** 1
Andrewes, C. H. (1949) *Proc. R. Soc. Med.* **42,** 519
Andrewes, C. H., Laidlaw, P. P. & Smith, W. (1934) *Lancet*, **2,** 859
Bozzo, A. (1952) *Bull. World Hlth Org.* **5,** 149
Bradley, W. H. (1951) *Proc. R. Soc. Med.* **44,** 789
Brans, L. M. (1952) *Studies on the antigenic composition of influenza virus B strains.* Stenfert Kroese, Leiden
Burnet, F. M. (1940) *Aust. J. exp. Biol. med. Sci.* **18,** 353
Burnet, F. M. (1951) *Proc. roy. Soc.* B, **138,** 47
Burnet, F. M. & Bull, D. R. (1943) *Aust. J. exp. Biol. med. Sci.* **21,** 55
Burnet, F. M. & Clark, E. (1942) *Influenza.* Macmillan, Melbourne
Burnet, F. M., Fraser, K. B. & Lind, P. E. (1953) *Nature, Lond.* **171,** 163
Chu, C. M. (1951) *J. gen. Microbiol.* **5,** 739
Chu, C. M., Andrewes, C. H. & Gledhill, A. W. (1950) *Bull. World Hlth Org.* **3,** 187
Chu, C. M., Dawson, I. M. & Elford, W. J. (1949) *Lancet*, **1,** 602
Davenport, F. M. & Francis, T., jr (1951) *J. exp. Med.* **93,** 129
Dawson, I. M. & Elford, W. J. (1949) *Nature, Lond.* **163,** 63
Elford, W. J., Andrewes, C. H. & Tang, F. F. (1936) *Brit. J. exp. Path.* **17,** 51
Fairbrother, R. W. & Hoyle, L. (1937) *J. Path. Bact.* **44,** 213
Fazekas de St Groth, S. (1950) *Lancet*, **1,** 1101
Fazekas de St Groth, S., Donnelley, M. & Graham, D. M. (1951) *Aust. J. exp. Biol. med. Sci.* **29,** 323
Francis, T., jr (1934) *Science*, **80,** 457
Francis, T., jr (1940) *Science*, **92,** 405
Francis, T., jr (1941) *Problems and trends in virus research.* University of Pennsylvania Press, Philadelphia
Francis, T., jr & Magill, T. P. (1935) *Science*, **82,** 353
Francis, T., jr & Magill, T. P. (1936) *Proc. Soc. exp. Biol., N.Y.* **33,** 604
Francis, T., jr, Pearson, H. E., Sullivan, E. R. & Brown, P. M. (1943) *Amer. J. Hyg.* **37,** 294
Francis, T., jr, Quilligan, J. J., jr & Minuse, E. (1950) *Science*, **112,** 495
Henle, W. (1950) *J. Immunol.* **64,** 203
Hennessen, W. A. (1952) *Bull. World Hlth Org.* **6,** 481
Hilleman, M. R. (1952) *Fed. Proc.* **11,** 798
Hilleman, M. R. & Horsfall, F. L., jr (1952) *J. Immunol.* **69,** 343
Hirst, G. K. (1941) *Science*, **94,** 22
Hirst, G. K. (1942) *J. exp. Med.* **76,** 195
Hirst, G. K. (1943) *J. exp. Med.* **78,** 407
Hirst, G. K. (1947) *J. exp. Med.* **86,** 357
Hirst, G. K. (1952) *J. exp. Med.* **96,** 589
Hoyle, L. (1944) *Mon. Bull. Minist. Hlth Lab. Serv.* **3,** 58
Hoyle, L. & Fairbrother, R. W. (1937) *J. Hyg., Camb.* **37,** 512
Isaacs, A. (1953) *Lancet*, **1,** 676
Isaacs, A. & Andrewes, C. H. (1951) *Brit. med. J.* **2,** 921
Isaacs, A., Edney, M., Donnelley, M. & Ingram, M. W. (1950) *Lancet*, **1,** 64
Isaacs, A., Gledhill, A. W. & Andrewes, C. H. (1952) *Bull. World Hlth Org.* **6,** 287
Isaacs, A. & Sampaio, A. A. C. (1953) *J. Path. Bact.* **65,** 613
Laidlaw, P. P. (1935) *Lancet*, **1,** 1118
McClelland, L. & Hare, R. (1941) *Canad. publ. Hlth J.* **32,** 530
Magill, T. P. (1940) *Proc. Soc. exp. Biol., N.Y.* **45,** 162
Magill, T. P. & Francis, T., jr (1936) *Proc. Soc. exp. Biol., N.Y.* **35,** 463
Magill, T. P. & Francis, T., jr (1938) *Brit. J. exp. Path.* **19,** 273
Magill, T. P. & Jotz, A. C. (1952) *J. Bact.* **64,** 619
Mulder, J., Brans, L. M. & Hers, J. F. P. (1952) *Leeuwenhoek ned. Tijdschr.* **18,** 131
Nigg, C., Crowley, J. H. & Wilson, D. E. (1940) *Science*, **91,** 603
Pettit, H., Mudd, S. & Pepper, D. S. (1936) *J. Amer. med. Ass.* **106,** 890
Salk, J. E., Bailey, M. L. & Laurent, A. M. (1952) *Amer. J. Hyg.* **55,** 439
Salk, J. E., Laurent, A. M. & Bailey, M. L. (1951) *Amer. J. publ. Hlth*, **41,** 669
Shope, R. E. (1931) *J. exp. Med.* **54,** 349
Shope, R. E. (1936) *J. exp. Med.* **63,** 669
Shope, R. E. (1943) Messenger Lecture in : Rivers *et al. Virus diseases*, p. 85. Cornell University Press, New York
Smith, W. (1935) *Brit. J. exp. Path.* **16,** 508
Smith, W. (1936) *Lancet*, **2,** 1256
Smith, W. & Andrewes, C. H. (1938) *Brit. J. exp. Path.* **19,** 293
Smith, W., Andrewes, C. H. & Laidlaw, P. P. (1933) *Lancet*, **2,** 66
Smith, W., Andrewes, C. H. & Laidlaw, P. P. (1935) *Brit. J. exp. Path.* **16,** 291
Stuart-Harris, C. H. (1939) *Lancet*, **1,** 487
Stuart-Harris, C. H., Andrewes, C. H. & Smith, W. (1938) *Spec. Rep. Ser. med. Res. Coun., Lond.* No. 228
Taylor, R. M. (1949) *Amer. J. publ. Hlth*, **39,** 171
USA Army Commission on Influenza (1944) *J. Amer. med. Ass.* **124,** 982
Veen, J. van der & Mulder, J. (1950) *Studies on the antigenic composition of human influenza virus A strains with the aid of the haemagglutination inhibition technique.* Stenfert Kroese, Leiden
Wang, C. (1948) *J. exp. Med.* **88,** 515

MUMPS

E. T. C. SPOONER M.D. M.R.C.S. L.R.C.P.

London School of Hygiene and Tropical Medicine
University of London

1 Susceptibility of experimental animals
2 Growth of the virus in the developing hen-egg
3 Physical properties of the virus
4 Complement fixation
5 Haemagglutination
6 Haemagglutinin inhibition
7 Strain differences
8 Preservation of virus
9 Relationship to other viruses
10 Immunity to mumps: vaccination
References

Mumps is a familiar disease " from Iceland to Egypt, from Alaska to Polynesia "—in all lands. It is a disease chiefly of children and young adults, but it rarely causes death or permanent damage, and is therefore not usually reckoned very serious. Consequently, statistical information about its prevalence is, in most countries, defective.

Such information as there is rests on interrogation of population samples (Logan, 1951). Occasionally, however, mumps may prove troublesome. Gordon (1940) quotes Haven Emerson that mumps was the most important " wastage disease " in the American Expeditionary Force in France in 1918. It has produced severe epidemics in garrisons, ships, schools and other institutions in which children or young adults congregate.

The clinical effects of mumps, which include the usual parotitis, and also orchitis, meningo-encephalitis, pancreatitis and other, rarer, localized lesions, have remained very constant, apparently since the days of Hippocrates.

1. Susceptibility of Experimental Animals

Nicolle & Conseil (1913) produced mumps in the monkey by intraperitoneal injection of material aspirated from the parotids of children suffering from the disease, and Gordon in 1914 (see Gordon, 1930), by intracerebral injection, produced a meningo-encephalitis with, in one animal, a parotitis. In 1934, Johnson and Goodpasture infected rhesus monkeys by inoculation into the parotid duct, and proved that the causal agent readily passed bacteriological filters (Johnson & Goodpasture, 1935). Findlay & Clarke (1934) confirmed this work in the same year.

Enders and his associates (Enders, Kane, Cohen & Levens, 1945; Enders, Cohen & Kane, 1945; Kane & Enders, 1945) used experimentally infected monkeys in the first intensive laboratory study of the disease. They prepared complement-fixing antigens and skin test reagents from infected monkey parotids, and with these studied human immunity to mumps.

No animal other than monkeys has been shown to be susceptible except the developing chick and, more recently, the suckling mouse (Kilham & Murphy, 1952).

2. Growth of the Virus in the Developing Hen-Egg

A great impetus was given to the study of mumps by the discovery of Habel (1945) that the virus can be propagated in the embryo chick. Rocchi, in 1943, claimed to have grown the virus on the chorio-allantoic membrane and to have produced parotitis in monkeys with fourth and sixth egg-passage material, but the chorio-allantoic membrane lesions which he depicts were probably due to something other than the mumps virus.

Habel (1945) infected eggs from monkey parotid gland material by inoculation into amniotic, allantoic or yolk sacs; yolk-sac inoculation, in his hands, gave the best results. Levens & Enders (1945) confirmed Habel's discovery.

Beveridge, Lind & Anderson (1946) isolated the virus from human saliva, by direct inoculation of eggs, in three out of eight cases. Many others have since isolated the virus in eggs (Leymaster & Ward, 1947; Hook, Poole & Friedewald, 1949; Wolff, 1950; Hannoun, 1951; Spooner, 1952; and others). Kilham (1948) isolated the virus from the blood of a patient, and it has been obtained by similar techniques from the cerebrospinal fluid of patients suffering from meningo-encephalitis caused by mumps (Henle & McDougall, 1947; Klone, 1952).

The Australian workers (Beveridge *et al.* 1946; Beveridge & Lind, 1946, 1947; Burnet, 1946) have made an intensive study of the behaviour of the mumps virus in the egg, in the course of which they have emphasized the resemblances and the differences between this virus and the viruses of influenza and Newcastle disease. Habel used the complement-fixation reaction to detect virus in inoculated eggs; haemagglutination (Levens & Enders, 1945; Beveridge & Lind, 1946) offers a quicker and easier method. Inoculated into the yolk sac, the virus appears therein and also in the amniotic fluid. Inoculated into the amniotic sac, it appears in the amniotic and occasionally in the allantoic fluid. Multiplication in the allantoic sac usually occurs only after adaptation to grow in eggs.

Compared with influenza virus, mumps virus multiplies slowly in the egg. Best yields are obtained (Beveridge & Lind, 1947) from seven- or eight-day-old eggs incubated for four or five days at 33°–36° C. after inoculation. Most workers have found that very little virus, as judged by complement fixation or haemagglutination, is produced in the first egg passages. Usually, of a dozen or so eggs inoculated with washings from the mouth, only one or two eggs will yield fluids showing either reaction, and then only to low titre; but after two or three passages, some adaptation to the egg usually happens.

It is uncertain whether " wild " virus will grow better in yolk sac or in amniotic sac; when trying to isolate a strain from mouth washings it is perhaps well to use both routes in the same or different eggs. The allantoic sac is not readily attacked by recently isolated virus. The Enders strain (Weller & Enders, 1948) will, after many egg passages, infect eggs regularly and to high titre when inoculated into the allantoic sac, but other strains (Lundbäck, 1949; Spooner, 1952) seem more difficult to adapt and may remain very erratic in their behaviour in the allantoic sac for many passages. Greenbury (unpublished) has found that, in about 50% of eggs inoculated allantoically with high concentration of a poorly adapted virus (Spooner, 1952), soluble antigen may develop in the chorio-allantoic membrane without detectable haemagglutinin appearing in the allantoic fluid.

3. Physical Properties of the Virus

Elford, Chu, Dawson, Dudgeon, Fulton & Smiles (1948) made a detailed comparison of the physical properties of the viruses of mumps, Newcastle disease and fowl plague. They studied the Enders strain of mumps, and found that gradocol filtration and centrifugation measurements gave it a particle diameter of about 100–150 mμ. Electron micrographs of purified virus indicated a more irregular size—from 85 mμ. to 200 mμ.—than is found with other viruses of the group. The particles appeared in electron micrographs as rather dome-shaped bodies, except that the larger particles were flattened. A strain isolated in England appeared similar, in an electron micrograph, to that of the American strain (Elford, personal communication).

The pH range of stability of infectivity was about 6.5–8.5.

4. Complement Fixation

Enders & Cohen (1942) showed that the sera of patients convalescent from mumps would fix complement specifically with antigens derived from infected monkey parotid. More recently, antigens prepared from virus-infected eggs have been used. Henle, Henle & Harris (1947) separated two complement-fixing antigens. One, the V or "virus" antigen, is firmly associated with the virus particle; the other, the S or "soluble" antigen, can be separated from virus particles by centrifugation. The V antigen is present abundantly in infected egg fluids; and the S antigen can be obtained in larger quantities from infected tissues, although both antigens are present in both situations. Antibodies to the two antigens can be separated by absorption. In man, S antibodies usually develop a day or two earlier than V antibodies and fall away more quickly (Henle, Harris & Henle, 1948). Study of titres to the two antigens may be of value in diagnosis of infection with the mumps virus (MacCallum, 1951). For diagnostic tests, the two antigens may be used separately, but for most routine purposes it suffices to use a mixed antigen consisting of untreated allantoic fluid, or a concentrate made from it (Sohier & Trimberger-Esser, 1952).

5. Haemagglutination

Levens & Enders (1945) discovered that the mumps virus in infected egg fluids will, like the influenza virus and Newcastle disease virus (NDV), agglutinate the red cells of fowls and mammals. This reaction has been studied by many workers, in particular by Lind (1948), Beveridge & Lind (1946) and others of the Burnet school, and by Robbins, Kilham, Levens & Enders (1949) and Sohier & Esser-Trimberger (1952). The Australian workers have shown that the mechanism of agglutination of the mumps virus is similar to that of the viruses of influenza and Newcastle disease. The virus particles attach themselves to the red cell, or to the red-cell stroma, by receptors on the cell. These receptors can be destroyed by the "receptor-destroying enzyme" of *Vibrio cholerae*, or, in suitable circumstances, by similar activity of the virus itself. At low temperatures, virus and cell unite, to produce haemagglutination. At higher temperatures, the virus enzyme breaks the bond, and virus is eluted. Cells from which virus has been eluted are no longer susceptible to haemagglutination by the same virus. Mumps virus (at 37° C.) will destroy the receptors by which it is itself attached to the red cell, but will leave others which can hold NDV or the influenza viruses. NDV and the influenza viruses will, however, remove their own and also the mumps receptors. This relationship is expressed in Burnet's conception of a "receptor gradient" (Burnet, McCrea & Stone, 1946), mumps falling at one end of the gradient.

Mumps haemagglutination may be measured by any of the usual techniques, including the photoelectric method of Belyavin, Westwood, Please & Wilson Smith (1951), which allows extrapolation of titre between tubes of a series.

Elution of mumps virus occurs very readily at 37° C., some batches of red cells releasing a better yield of virus than others. Purification of the virus by elution is complicated by the fact that the virus causes haemolysis, but sheep cell stromata may be used for the purpose; it is, however, unusual to be able to recover more than half of the haemagglutinating virus, and the method is irregular. Elford *et al.* (1948) found purification by elution less satisfactory than with influenza. Forster & Carson (1949) preferred to concentrate the virus by a methanol precipitation method.

"Indicator" virus cannot be prepared from mumps virus as it can, for example, from influenza B virus, by heating, because the mumps haemagglutinin itself is about as heat-labile as the eluting enzyme.

6. Haemagglutinin Inhibition

Specific antibody inhibits haemagglutination. Haemagglutinin inhibition titrations can be used for the serological diagnosis of mumps infection (Beveridge & Lind, 1946). Aikawa & Meiklejohn (1949) compared the method with complement fixation and found that haemagglutinin inhibitor titres rose to their maximum rather more quickly. They began to rise at about the same stage of the illness as did the complement-fixing antibodies. One reason why haemagglutinin inhibition tests are not popular for diagnostic purposes is that non-specific inhibitors occur in some sera, as they do for influenza.

Burnet (1946) devised another diagnostic test, depending on the fact that specific antisera will agglutinate red cells which have been exposed to mumps virus in amniotic fluid, and from which virus has subsequently eluted to the point of restoring the suspension stability of the cells. Virus in allantoic fluid is unsatisfactory. Others have found this test irregular, and perhaps too sensitive for routine use.

7. Strain Differences

To date, no antigenic differences have been discovered between the rather small number of strains that have been compared with one another. Strains may differ in their virulence for monkeys and in their behaviour in eggs. Ginsberg & Horsfall (1949) produced a strain which was abnormally resistant to the inhibitory effect of Friedländer's bacillus polysaccharide. Spooner (1952) has found that the Enders strain differs from a British strain in that its haemagglutinin is slightly more heat-resistant.

8. Preservation of Virus

Mumps virus can be freeze-dried. A satisfactory method is to dilute infected egg fluid with an equal volume of 14% sucrose. Such preparations, freeze-dried and held at − 20° C., remain infective for at least four years. The virus, in sealed ampoules, may also be preserved at − 65° C. At + 1° C., egg fluids remain infective for a week or two. Bennett

& Kempf (1952) have recently made the interesting observation that the complement-fixing antigen is much more stable at + 5° C. than at sub-freezing temperatures.

9. Relationship to Other Viruses

Particle size and haemagglutination properties place the mumps virus alongside the influenza and Newcastle disease viruses. It differs from these antigenically, and in being more slow-growing and less regular in particle size. It also produces a haemolysin (Morgan, Enders & Wagley, 1948).

Several workers have found an antihaemagglutinin for Newcastle disease in sera of patients convalescent from mumps, and have suggested some antigenic relationship between the two viruses on this account. However, careful studies by Evans (1951), and by Wenner, Jenson & Monley (1952) and by Wenner, Monley & Jenson (1952), have failed to reveal any antigenic relationship demonstrable by experiments on monkeys and rabbits. Wenner and his associates conclude that the apparent cross-reaction is due to non-specific inhibitors; Evans suggests that infection with mumps may, in humans, produce an antibody which reacts also with NDV. All strains of NDV do not behave alike in this respect.

10. Immunity to Mumps: Vaccination

An attack of mumps probably produces a good and lasting immunity, for second attacks are unusual. Mumps epidemics, perhaps because of the long incubation period of the disease, are inclined to last for a long time. Often, only a small percentage of the presumed susceptibles (not previously attacked) who are exposed to risk of infection contract the disease (Medical Research Council, 1938).

Antibodies detectable by complement-fixation, haemagglutination or neutralization tests develop in convalescent men and monkeys. Greenbury (unpublished) has found that antibodies to both S and V antigens develop regularly in rats and mice, but not in guinea-pigs, inoculated intranasally, although these animals do not become infected in the ordinary sense.

The first attempts to produce artificial immunity in man were made by Enders and his associates, using an antigen made from monkey parotids. In 1946, Enders, Levens & Maris found that 15 to 25 passages in eggs rendered their strain avirulent for monkeys. Beveridge & Lind (1947) vaccinated children with a formolized egg-fluid vaccine, and were able to show that antibody formation followed; they had no opportunity to test the efficacy of the method against natural infection. The most extensive tests on vaccines inactivated with formalin or ultra-violet light so far made have been those of Habel (1951); but more work will be needed to establish the most satisfactory virus. In any case, the value of vaccines will probably be limited to control of the disease in specially susceptible communities; it is unlikely that widespread vaccination would justify its cost.

REFERENCES

Aikawa, S. K. & Meiklejohn, G. (1949) *J. Immunol.* **62,** 261
Belyavin, G., Westwood, J. C. N., Please, N. W. & Smith, W. (1951) *J. gen. Microbiol.* **5,** 546
Bennett, A. G. & Kempf, J. E. (1952) *Proc. Soc. exp. Biol., N.Y.* **81,** 400
Beveridge, W. I. B. & Lind, P. E. (1946) *Aust. J. exp. Biol. med. Sci.* **24,** 127
Beveridge, W. I. B. & Lind, P. E. (1947) *Aust. J. exp. Biol. med. Sci.* **25,** 337
Beveridge, W. I. B., Lind, P. E. & Anderson, S. G. (1946) *Aust. J. exp. Biol. med. Sci.* **24,** 15
Burnet, F. M. (1946) *Brit. J. exp. Path.* **27,** 244
Burnet, F. M., McCrea, J. F. & Stone, J. D. (1946) *Brit. J. exp. Path.* **27,** 228
Elford, W. J., Chu, C. M., Dawson, I. M., Dudgeon, J. A., Fulton, F. & Smiles, J. (1948) *Brit. J. exp. Path.* **29,** 590
Enders, J. F. & Cohen, S. (1942) *Proc. Soc. exp. Biol., N.Y.* **50,** 180
Enders, J. F., Cohen, S. & Kane, L. W. (1945) *J. exp. Med.* **81,** 119
Enders, J. F., Kane, L. W., Cohen, S. & Levens, J. H. (1945) *J. exp. Med.* **81,** 93
Enders, J. F., Levens, J. H. & Maris, E. P. (1946) *J. Immunol.* **54,** 283
Evans, A. S. (1951) *J. Immunol.* **67,** 529
Findlay, G. M. & Clarke, L. P. (1934) *Brit. J. exp. Path.* **15,** 309
Forster, G. F. & Carson, E. (1949) *J. infect. Dis.* **85,** 62
Ginsberg, H. S. & Horsfall, F. L., jr (1949) *J. exp. Med.* **90,** 393
Gordon, J. E. (1940) *Amer. J. med. Sci.* **200,** 412
Gordon, M. H. (1930) In: Medical Research Council. *A system of bacteriology in relation to medicine,* vol. 7, p. 273. His Majesty's Stationery Office, London
Habel, K. (1945) *Publ. Hlth Rep., Wash.* **60,** 201
Habel, K. (1951) *Amer. J. Hyg.* **54,** 295
Hannoun, C. (1951) *Ann. Inst. Pasteur,* **80,** 175
Henle, G., Harris, S. & Henle, W. (1948) *J. exp. Med.* **88,** 133
Henle, G., Henle, W. & Harris, S. (1947) *Proc. Soc. exp. Biol., N.Y.* **64,** 290
Henle, G. & McDougall, C. L. (1947) *Proc. Soc. exp. Biol., N.Y.* **66,** 209
Hook, E. W., jr, Poole, S. O. & Friedewald, W. F. (1949) *J. infect. Dis.* **84,** 230
Johnson, C. D. & Goodpasture, E. W. (1935) *Amer. J. Hyg.* **21,** 46
Kane, L. W. & Enders, J. F. (1945) *J. exp. Med.* **81,** 137
Kilham, L. (1948) *Proc. Soc. exp. Biol., N.Y.* **69,** 99
Kilham, L. & Murphy, H. W. (1952) *Proc. Soc. exp. Biol., N.Y.* **80,** 495
Klone, W. (1952) *Klin. Wschr.* **30,** 782
Levens, J. H. & Enders, J. F. (1945) *Science,* **102,** 117
Leymaster, G. R. & Ward, T. G. (1947) *J. clin. Invest.* **26,** 1188
Lind, P. E. (1948) *Aust. J. exp. Biol. med. Sci.* **26,** 93
Logan, W. P. D. (1951) *Mon. Bull. Minist. Hlth Lab. Serv.* **10,** 136
Lundbäck, H. (1949) *Brit. J. exp. Path.* **30,** 221
MacCallum, F. O. (1951) *Mon. Bull. Minist. Hlth Lab. Serv.* **10,** 267
Medical Research Council (1938) *Spec. Rep. Ser. med. Res. Coun., Lond.* No. 227
Morgan, H. R., Enders, J. F. & Wagley, P. F. (1948) *J. exp. Med.* **88,** 503
Nicolle, C. & Conseil, E. (1913) *C. R. Soc. Biol., Paris,* **92,** 1370
Robbins, F. C., Kilham, L., Levens, J. H. & Enders, J. F. (1949) *J. Immunol.* **61,** 235
Rocchi, F. (1943) *Arch. ges. Virusforsch.* **2,** 499
Sohier, R. & Esser-Trimberger, I. (1952) *Ann. Inst. Pasteur,* **82,** 696
Sohier, R. & Trimberger-Esser, I. (1952) *Ann. Inst. Pasteur,* **82,** 168
Spooner, E. T. C. (1952) *Mon. Bull. Minist. Hlth Lab. Serv.* **11,** 94
Weller, T. H. & Enders, J. F. (1948) *Proc. Soc. exp. Biol., N.Y.* **69,** 124
Wenner, H. A., Jenson, M. H. & Monley, A. (1952) *J. Immunol.* **68,** 343
Wenner, H. A., Monley, A. & Jenson, M. H. (1952) *J. Immunol.* **68,** 357
Wolff, H. L. (1950) *Ned. Tijdschr. Geneesk.* **94,** 1112

YELLOW FEVER

A Problem in Epidemiology

G. W. A. DICK B.Sc. M.D. M.R.C.P.Ed. M.P.H.

National Institute for Medical Research
Mill Hill, London

1 Introduction
2 Yellow fever in South America
3 Yellow fever in Africa
4 Control of yellow fever
5 Viruses isolated during yellow fever investigations
References

1. Introduction

The recovery by Dr A. F. Mahaffy of a strain of yellow fever virus from an African called Asibi and the subsequent propagation of the now famous Asibi strain in rhesus monkeys by Stokes, Bauer & Hudson (1928a) mark the beginning of the modern period of research on yellow fever. Although Reed & Carroll (1902) had already shown that the disease could be transmitted with Berkefeld filtrates of diluted infective human serum, this finding of a susceptible animal led rapidly to the characterization of the infective agent as a virus which could be recovered from filtrates of infective blood and of infected mosquitoes (Stokes, Bauer & Hudson, 1928b; Bauer & Mahaffy, 1930a) and was of a particle size of 18–27 mμ. (Findlay & Broom, 1933). Other strains of the virus were soon isolated and Sawyer, Kitchen, Frobisher & Lloyd (1930) were able to demonstrate by cross-immunity tests in rhesus monkeys that the African and the South American yellow fever were the same disease. Another great advance was made by Theiler (1930), who adapted the French strain of virus (Mathis, Sellards & Laigret, 1928) to mice, and developed a mouse neutralization test (Theiler, 1931) which, as modified by Sawyer & Lloyd (1931), made extensive immunological and epidemiological studies possible. Although it had long been recognized that there are great variations in the severity of the infection, laboratory investigations established that the classical disease, associated with jaundice, haemorrhage and black vomit, is no more typical of infection with yellow fever virus than is paralysis now regarded as a usual result of infection with poliomyelitis and that the great bulk of infections are (i) subclinical (which may be diagnosed only by the recovery of virus or by the demonstration of antibodies), or (ii) of a mild nature. The mild cases, which are probably the most common type of clinical infection, are unlikely to be diagnosed except during an epidemic or during search for a case in an endemic area. They are associated with pyrexia and symptoms of fever and headache lasting for two to three days. The clinical diagnosis in these cases may be suggested by albuminuria, and by bradycardia in convalescence.

The early epidemiological studies which had been made had indicated that *Aedes aegypti* was the only vector of the disease and apparent confirmation of this was obtained in North, Central and South America from the rapid disappearance of yellow fever when *A. aegypti* control was put into effect. Indeed, so successful were the control methods in the Americas that by 1925 the only area of infection was a small focus in Brazil. The epidemiology of the disease seemed simple and was dependent on an *A. aegypti*-man cycle, which was maintained by adequate numbers of both *A. aegypti* and susceptible humans. However, Bauer (1928) and Philip (1929, 1930a, 1930b) showed that in Africa, transmission of the disease could be effected experimentally by mosquitoes other than *A. aegypti*, and Bauer & Mahaffy (1930b) demonstrated that several species of African monkeys were susceptible to infection in so far as they circulated virus for several days after inoculation. These findings developed the hypothesis that a cycle of infection might exist involving forest mosquitoes and monkeys, for Balfour (1914) had suggested many years previously that monkeys were the reservoir of yellow fever. Meanwhile, in South America, Davis & Shannon (1929) had reported the successful transmission of yellow fever by mosquitoes other than *A. aegypti*, and Davis (1930a, 1930b) found cebus monkeys and marmosets were, like the African monkeys, susceptible to infection. These observations shook the concept of the simple *A. aegypti*-man cycle, but the possibility of a cycle of infection existing in the absence of *A. aegypti* was not convincingly established till Soper, Penna, Cardoso, Serafim, Frobisher & Pinheiro (1933) presented the results of epidemiological studies made in Brazil, which indicated that human beings had acquired infection in the absence of *A. aegypti*, but as a result of contact with forest, and for this form of infection Soper (1935) used the term "jungle yellow fever", although "sylvan yellow fever" is, in the light of present knowledge, perhaps a more descriptive term.

2. Yellow Fever in South America

It soon became apparent that in South America there were two cycles of infection: a man-mosquito-man cycle of which *A. aegypti* was the unique vector, and a cycle or cycles in which contact with forest was a sine qua non of infection. In the Americas *A. aegypti* is a highly domestic mosquito and as might be expected all age-groups and both sexes were equally infected when the virus was first introduced into a community. On reintroduction of the virus, those individuals who have previously encountered it will not be affected, for after infection a durable immunity develops. In general it might be said that yellow fever transmitted by *A. aegypti* was a disease of "stay-at-homes", of travellers entering, or of children being born into, an infected area. The disease occurred epidemically or was endemic in both urban and rural areas.

Early investigations of the forest cycle by Soper (1936a, 1936b) showed that monkeys from areas where yellow fever was present had neutralizing antibody in their sera, while those from areas where yellow fever was absent had no immunity, as evidenced by neutralization tests. More intensive studies of this type (Taylor & Fonseca da Cunha, 1946; Laemmert, Ferreira & Taylor, 1946; Kumm & Laemmert, 1950) revealed that naturally acquired immunity is present in all species of South American monkeys that have been tested in significant numbers. The susceptibility of many of the monkeys has been tested in the laboratory by inoculation of virus by syringe or by the bite of infected mosquitoes and the possibility of a cyclic transmission from monkey to monkey by mosquito bite has been demonstrated. It is clear from all these studies that monkeys are important hosts of the virus in

South America. The part which animals other than primates may play in a forest cycle is much less certain. Of these most attention has been directed to *Marsupialia*, particularly by Bugher, Boshell-Manrique, Roca-García & Gilmore (1941). Bugher (1951) would appear to attach more importance to their epidemiological significance than do other investigators. There appears to be considerable variation in the susceptibility and immunological response of the same species in different geographic regions and in their response to different strains of virus. While some of the opossums have been found in nature to have sera which neutralized yellow fever virus, the specificity of such positive tests has been questioned (Bates, 1944; Laemmert, 1946; Koprowski, 1946; Laemmert, Ferreira & Taylor, 1946). The latter investigator (Taylor, 1951) is of the opinion that of the opossums only the genus *Metachirus* and some species of the genus *Marmosa* have been shown to be "epidemiologically susceptible" under laboratory conditions. Except for the primates and perhaps the marsupials (which require further study), it does not appear likely that other animals found to be susceptible in the laboratory (such as some of the rodents and carnivores) fit in to the epidemiological picture.

Of a number of forest mosquitoes found capable of transmitting the virus experimentally, it seems that the most likely vectors in South America are species of the genus *Haemagogus* and *Aedes leucocelaenus*. Not only are these species efficient laboratory vectors, but in 1938 Shannon, Whitman & Franca successfully infected monkeys by exposing them to the bite of wild caught mosquitoes of the species *H. capricornii* and *A. leucocelaenus* and, since then, virus has been obtained eighteen times from species of *Haemagogus* (*capricornii, spegazzinii* and *spegazzinii falco*) and three times from *A. leucocelaenus*. Many things point to *Haemagogus* spp. being the most important vectors, for, although the different species have different distributions in South and Central America, in general this genus covers the entire region over which the forest cycle of yellow fever is known to exist. The adults of the genus are found in the greatest numbers in the forest canopy (Bugher, Boshell-Manrique, Roca-García & Osorno-Mesa, 1944). They are essentially tree-hole breeders and daylight feeders, being most active at midday, and occasionally descend to ground level at the forest edge and bite human beings. While *A. leucocelaenus* is of somewhat similar habitat and is widely distributed in South America (Kumm & Cerqueira, 1951), it shows no preference for the forest canopy and in some forests in which yellow fever is present this mosquito is rare (Laemmert *et al.* 1946; Kumm, 1950). It would appear then that the virus may be maintained by a monkey-mosquito cycle in which *Haemagogus* would seem to be the most important vector. The midday biting activity of the *Haemagogus* makes resting monkeys an easy prey and the persistence of *Haemagogus* throughout the dry season permits the mosquito to act as the true reservoir of the virus (Bugher *et al.* 1944). Man is but accidentally involved in the cycle by being bitten by an infected mosquito which has come down to ground level or else, for example, during tree felling when the mosquitoes of the canopy are brought down with the tree. It appears that the forest cycle is enzootic in some areas and from these enzootic foci epizootics spread to other regions. Taylor (1951) likened these epizootics to the pseudopodia of an amoeba "protruding outward and then retracting or rather vanishing as the available supply of susceptible hosts is used up". In the enzootic zones such as the great rain forest of the Amazon-Orinoco watershed and the Ilhéus area in Brazil, the Magdalena basin and the contiguous forests in Venezuela and the Guianas (Boshell-Manrique, 1948), the cycle of infection is apparently maintained by a series of small wandering epizootics. The disease among human beings in these areas is also endemic, as evidenced by antibody surveys and the finding of a few positive viscerotomy specimens every year. A remarkable feature of the epizootic excursions is their rapidity of spread, as evidenced by the finding of monkeys which have died of yellow fever and the simultaneous or slightly later appearance of the disease in human beings. The rate of some of these epizootic waves has been calculated. Thus in November and December 1948 a wave of sylvan yellow fever from an enzootic focus in eastern Panama moved westward following the Atlantic rain forest and reached north Costa Rica in July 1951. Early in October 1951 it crossed to the Pacific side and re-entered western Panama. The velocity of the wave was 12–15 miles per month (Clark, 1952; Elton, 1952a, 1952b, 1952c). How this wave will continue is uncertain but it is now reaching up towards Nicaragua and is calculated to reach Mexico in about three or four years. Similar waves have been described by Taylor (1951) in Brazil —one in 1934 being spread by successive yearly steps for about 1,000 miles. These excursions go in halts and bursts, being limited in their activity to the warm rainy season when mosquito activity is greatest, and disappear and reappear further on with the onset of the rains and warm weather. In order to locate the spread and activity of the virus several novel methods of study were introduced, such as the placing of normal monkeys in cages in the forest canopy (sentinel monkeys) and the tagging and recapturing of cebus monkeys and the periodic testing of their sera for antibody. Neither of these methods indicated virus activity and most information on the spread of virus has been obtained from viscerotomy and serological studies in human beings. The epidemics of jungle or preferably sylvan yellow fever show no consistent pattern in their periodicity (Kumm, 1952) and the way in which they spread is not yet completely understood. Since monkeys in South America do not migrate from one forest patch to another (Causey, Laemmert & Hayes, 1948) they cannot be responsible for the transport of the virus beyond a very limited area. Recent work by Wellington (1945) indicates that small insects may be transported long distances by convection currents and wind, and Causey, Kumm & Laemmert (1950) were able to recapture *H. spegazzinii* labelled with bronzing powder more than seven miles from the point of release. It is necessary to postulate that the mosquitoes are carried much greater distances—of perhaps 30 to 40 miles—in order to explain the rate of spread of the epidemic waves (Taylor, 1951). As one area becomes infected with mosquitoes carrying virus, a further spread would depend on the availability of vectors; new supplies of mosquitoes would then be infected and, depending on the climate and environment, another jump could be made to a more distant focus. Since the incriminated vectors can survive the dry season, it is postulated that the virus is maintained in the mosquitoes which survive throughout that period. Further studies are required to establish whether this hypothesis will explain all the observed facts. The possibility that the virus is carried from forest to forest by birds with viraemia or carrying infected ectoparasites presents an entertaining hypothesis, but of this there is no evidence although experimental investigations along this line are far from being exhausted.

3. Yellow Fever in Africa

In Africa there appears to be a man-to-man cycle and a sylvan cycle of infection, but the study of yellow fever in Africa is beset with a number of obstacles. Thus while antibody studies (Beeuwkes & Mahaffy, 1934; Sawyer & Whitman, 1936; Mahaffy, Smithburn & Hughes, 1946; Smithburn, Goodner, Dick, Kitchen & Ross, 1949) have indicated a wide distribution of yellow fever in the peoples of Africa between Capricorn and Cancer, the difficulties of establishing a viscerotomy service (except to a certain extent in the Belgian Congo) and of rapidly and accurately obtaining and investigating reports of sickness and death have limited the knowledge of the behaviour of the disease in human beings. Other complications arise in that many Africans do not know when they were born, and whether or not they have been vaccinated against yellow fever, and further, with the migration of labourers (see Dick, 1953; Smithburn, Goodner, Dick, Kitchen & Ross, 1949), the finding of immune unvaccinated persons in any one area does not necessarily mean that they obtained their infection in the place where their blood samples were taken. All this, combined with the absence of vital statistics, makes accurate epidemiological studies difficult. Other difficulties arise from the fact that certain of the incriminated vectors in Africa show considerable variations in their ecology. In certain parts of West Africa *A. aegypti* bites human beings and is fairly domestic in its habits, in other parts it has been found breeding in tree holes and pools at considerable distances from the nearest human habitation (Dunn, 1927a, 1927b); while in some parts of Uganda it is essentially zoophilic, sylvan in habitat, and rarely found in towns and villages (Haddow, 1945a), in others it is domestic or semi-domestic and bites human beings (Gillett, 1951).

The complexity of the environments and lack of investigations in many places where yellow fever occurs in Africa make it impossible to present a general picture. In West Africa, where the disease is endemic and locally epidemic among human beings, it seems likely that *A. aegypti* is responsible for the transmission of yellow fever to man in the urban and village communities. Bugher is of the opinion that all human yellow fever in Nigeria from 1941 to 1951 can best be explained by a man-*A. aegypti* cycle (Whitman, 1951). It would seem that such was also the situation prior to 1951 (see Beeuwkes & Mahaffy, 1934) and as far as is known this view still holds. Virus was isolated from this species by Beeuwkes & Hayne (1931) and on three subsequent occasions in 1946 by Bugher and his colleagues (see Taylor, 1951). In Uganda there is no evidence that *A. aegypti* plays any part in the epidemiology of yellow fever and its place is taken by *A. simpsoni*. Virus was isolated from *A. simpsoni* taken in Bwamba, Uganda, in 1941 and this species was incriminated as the human vector of yellow fever in that area (Mahaffy, Smithburn, Jacobs & Gillett, 1942). This was the first indication that a mosquito other than *A. aegypti* was involved in transmitting the disease to human beings in Africa, although nine African species (including *A. simpsoni*) were at that time known to be capable of transmitting the virus in the laboratory (see Bauer, 1928; Philip, 1929, 1930a, 1930b; Kerr, 1932). (The number of known laboratory vectors is now twelve; see Lewis, Hughes & Mahaffy, 1942, and Bruce-Chwatt, 1950.) A second isolation from *A. simpsoni* was made in 1942 in Bwamba (Smithburn & Haddow, 1946). In that area this species bites human beings (Haddow, 1945b) and is semi-domestic in that it is found breeding in the axils of certain plants in the plantations around human dwellings (Haddow, 1945b, 1945c). In other parts of Uganda, however, *A. simpsoni* does not bite man (Gillett, 1951) and could not therefore take part in a man-mosquito cycle of infection.

Investigations of the most extensive epidemic that has ever been reported from Africa, and which broke out in 1941 in the Nuba Mountains of the Anglo-Egyptian Sudan, indicated that *A. aegypti* may have been the responsible vector (Kirk, 1941). On the other hand, Lewis (1943) believed that *A. vittatus* played a very important role in transmitting the virus in that area.

Because of the habitats of the human vectors of yellow fever it is not possible to eradicate the man-mosquito cycle in Africa, as is being done in South America, and since an apparent man-mosquito cycle may exist in an area where there is a forest cycle (which will next be discussed) it is well-nigh impossible to know whether the virus is maintained regionally endemic by a man-mosquito cycle or whether virus is being constantly introduced from the forests. This difficulty does not arise in South America, where it can now be assumed that all yellow fever is due to infection by forest mosquitoes.

Following the pioneer investigations into the susceptibility of West African monkeys to yellow fever by Bauer & Mahaffy (1930a), more extensive studies of the primates of central, East and West Africa were made by Hughes (1943), Smithburn & Haddow (1949) and by Bugher and his colleagues (Bugher, 1951). It may be concluded from all these investigations that (i) all the monkeys and apes that have been tested are susceptible to yellow fever in that virus multiplies after inoculation by syringe or by the bite of an infected mosquito and that the virus could be transmitted by mosquito bite to other monkeys, (ii) apart from slight pyrexia monkeys rarely show any signs of yellow fever infection and develop a solid immunity as a result of infection. It was further shown (Ross & Gillett, 1950) that yellow fever virus could be maintained by cyclic passage in monkeys and mosquitoes. Antibody surveys have indicated that monkeys in many of the forested areas of Africa are naturally immune (Findlay, Stefanopoulo, Davey & Mahaffy, 1936; Findlay & MacCallum, 1937; Haddow, Dick, Lumsden & Smithburn, 1951; Bugher, 1951). There is good evidence therefore that monkeys are important hosts of the virus, although it has not yet been possible to recover virus directly from naturally infected monkeys. (With the exception of grivet monkeys it is exceedingly difficult to trap monkeys in East and central Africa because of the abundance of available food.) It has also been shown that certain of the Lemuroidea (galagos, or bush-babies, and pottos) are susceptible (Smithburn, 1949; Bugher, 1951; Dick, unpublished) and high rates of antibody have been found in some of these species in the driest parts of Uganda and in Kenya (Haddow, 1952). It is possible that bush-babies rather than monkeys may be important hosts in certain areas where it seems unlikely that monkeys are extensively involved.

The possibility that animals other than primates may take part in the forest cycle in Africa has been investigated by Smithburn & Haddow (1949) and by Dick (1952a). It seems unlikely that any non-primates play an important role, if indeed any part, in the sylvan cycle. The most highly susceptible of all non-primates tested are hedgehogs, but they have not been implicated epidemiologically, although common in some areas.

A. africanus is the principal sylvan vector and host of yellow fever in Bwamba, Uganda (Haddow, Smithburn, Dick, Kitchen

& Lumsden, 1948; Smithburn, Haddow & Lumsden, 1949). It seems probable that such is the case also in other areas of lowland rain forest and forested areas in central Africa. This mosquito is found in greatest numbers in the forest canopy (Haddow, Gillett & Highton, 1947), it bites monkeys (Haddow & Dick, 1948) and shows a pronounced peak of biting activity just after dusk when all species of monkeys are resting or asleep. It also survives the dry season in the adult state in wooded and forested areas (Haddow & Mahaffy, 1949). *A. africanus* (and *A. luteocephalus* which may represent the extreme forms of the same species (Haddow & Mahaffy, 1949)) have a wide distribution across tropical Africa from the Anglo-Egyptian Sudan (Lewis, 1943) as far south as Bechuanaland (as found by De Meillon; see South African Institute for Medical Research, 1946). Whether this species alone or other sylvan vectors are involved in a forest cycle is not known. In order to explain the different immunity rates in partly terrestrial as compared with arboreal species of monkeys Haddow *et al.* (1951) have postulated that there may be a diurnal as well as the nocturnal vector (*A. africanus*). All attempts to isolate yellow fever virus in Uganda from many thousands of forest mosquitoes other than *A. africanus* have, however, been unsuccessful. It would seem therefore that in the lowland African rain forest a cycle of infection revolves around primates and *A. africanus*, similar to the primate-*Haemagogus* cycle in South and Central America.

How much further the analogy can be carried is uncertain. Although *A. africanus* is known to bite man in plantations and on the forest floor (Haddow, 1945b, 1945c; Gillett, 1951), because of its rarity at ground level in those areas that have been investigated it could not account for more than a few infections. (This is borne out by the finding of low rates of antibody in forest gangs and pygmies in Uganda.) This species may, however, be responsible for the sporadic infections of yellow fever which have occurred in people living in close contact with certain forests (Dick, 1953) in which yellow fever is enzootic in monkeys but where *A. simpsoni* does not bite man (Gillett, 1951). Thus, sylvan yellow fever, analogous to what is found in South America, may exist in certain parts of Africa: i.e., yellow fever contracted by the bite of a sylvan mosquito infected from a forest animal, not necessarily in the absence of *A. aegypti*, but in areas where *A. aegypti* has no importance as a vector. Indeed the human immunity rates found in many parts of central and East Africa can be explained better by postulating the existence of sylvan yellow fever than of a man-mosquito-man cycle.

The situation in Bwamba, Uganda, may be unique. In that area epidemic and endemic yellow fever occur (Mahaffy *et al.* 1946; Dick, 1950). The highest incidence of infection is found among people living in close proximity to the forest and there is a considerable amount of evidence which suggests that *A. simpsoni*—the human vector—probably obtains its infection from forest monkeys, particularly when they raid plantations (Smithburn & Haddow, 1946; Lumsden, 1951). On the other hand it also seems possible that in Bwamba the virus is introduced to the human population from time to time by a stray *A. africanus* from the forest; man-to-man transmission by *A. simpsoni* presumably then produces small localized epidemics. It is not known if other areas of Africa have an epidemiological picture such as has been postulated to explain the human epidemiology in Bwamba.

The complete geographical distribution of human yellow fever or of the forest cycle in Africa is not yet known. Nothing is known of the possible existence or spread of waves of infection such as exist in South and Central America. Although one isolated epizootic in monkeys has been recorded (Haddow *et al.* 1951), it is not known whether this represented the introduction of virus from elsewhere or was one of many little epizootics which together produce the enzootic picture in lowland rain forests. It is not known what cycle of infection exists in areas where high immunity rates have been found in bush-babies but not in monkeys. Little is known of the ecology of *A. africanus* or of *A. simpsoni* outside of Uganda, and it is not known how yellow fever persists in the drier areas where it seems unlikely that adequate numbers of *A. africanus* could survive a long season of drought. There is no evidence that the virus can be maintained by transovarial passage from one adult generation of mosquitoes to the next (Gillett, Ross, Dick, Haddow & Hewitt, 1950), and it remains to be seen whether the sylvan cycle in Africa is maintained by the primate-mosquito cycle or whether it is but the end-product of some less tenuous cycle which has not yet been discovered.

4. Control of Yellow Fever

Yellow fever transmitted by *A. aegypti* has been controlled in South America and around the airports and urban centres of many African towns by anti-mosquito measures. Control measures have also been applied in India and elsewhere, particularly around international airports. It is fortunate that yellow fever has not reached India; this may be due to the failure of Indian *A. aegypti* to maintain the virus. It would seem important that this should be investigated. However, since yellow fever has not entered India in the past, it seems unlikely to do so in the future.

The control of sylvan yellow fever in the Americas and of yellow fever in Africa is achieved mainly by vaccination. Tissue culture of the Asibi strain (Lloyd, Theiler & Ricci, 1936; Theiler & Smith, 1937a) induced a marked loss of viscerotropic and neurotropic affinities of the virus and resulted in the 17D strain which was successfully used for vaccination (Theiler & Smith, 1937b). The 17D vaccine as now used is prepared from developing chick embryos and the vaccine is inoculated subcutaneously. The efficacy of the vaccine has been clearly established. Immunity as evidenced by the demonstration of neutralizing antibodies lasts for at least nine years in a high proportion of vaccinated persons (Dick & Gee, 1952) but it is not certain if this applies to children vaccinated as babies. It is not known what level of maternal antibody is transmitted to the babies of vaccinated mothers. It seems possible, however, that if babies with maternal antibody from vaccinated mothers were vaccinated, passive immunity might prevent the active immunization of such babies. This might to a certain extent explain the apparent differences in the durability of immunity after vaccination in children as compared with older persons, recorded by Dick & Gee (1952) and by Fox & Cabral (1943).

French workers have employed a vaccine prepared from French neurotropic mouse-brain virus with or without the addition of vaccinia vaccine (Peltier, 1947, 1948), which is administered by scarification. Because of the dangers of encephalitis attending the use of this vaccine (see Macnamara, 1953) it has not been universally accepted, but the scarification technique has been used to administer 17D vaccine with excellent results (Hahn, 1951; Dick, 1952b) and the administration of a vaccine by scarification presents several advantages

for mass immunization. While Hahn (1951) obtained good results with a combined 17D and vaccinia vaccine, studies by Dick & Horgan (1952) indicated that the number of those who became immune to yellow fever after vaccination with a similar type of mixed vaccine which they prepared was not sufficiently high to merit the routine use of such a mixed vaccine without first making further studies of the optimum quantities of each virus required to effect adequate immunization against both smallpox and yellow fever.

5. Viruses Isolated during Yellow Fever Investigations

The techniques used in attempts to isolate yellow fever virus have yielded a number of hitherto unknown viruses. In Africa, West Nile (Smithburn, Hughes, Burke & Paul, 1940) and Bwamba fever viruses (Smithburn, Mahaffy & Paul, 1941) were isolated from human beings in Uganda and the former virus has been recovered from children in Egypt (Melnick, Paul, Riordan, Barnet, Goldblum & Zabin, 1951). Semliki Forest virus (Smithburn & Haddow, 1944), Bunyamwera (Smithburn, Haddow & Mahaffy, 1946), Ntaya (Smithburn & Haddow, 1951), Uganda S (Dick & Haddow, 1952) and Zika virus (Dick, Kitchen & Haddow, 1952) were isolated from mosquitoes. A strain of Zika virus was also recovered from the blood of a pyrexial "sentinel" monkey (Dick, Kitchen & Haddow, 1952) and has recently been isolated from a human being in West Africa (Virus Research Institute, Yaba, Nigeria: unpublished annual report, 1952). In addition, several strains of Mengo virus (Dick, Smithburn & Haddow, 1948) (now shown to be related to Columbia-SK and EMC viruses (Dick, 1949)) have been isolated from man, rhesus monkeys and mosquitoes. Antibody to all of these viruses has been found in sera of the residents of several places in Africa (see Dick, 1953) and to some of them in human sera taken in India (Smithburn, 1953). Bwamba fever and Zika viruses can cause a mild pyrexial disease, Mengo a mild encephalitis, and West Nile virus apparently produces a relatively silent viraemia in human beings; but nothing is known of the illnesses, if any, that the other agents cause. Antibodies to several of them have been found in monkeys, but practically nothing is known of their epidemiology. What part any of these viruses may play in the causation of sporadic or epidemic diseases in temperate climes has not been investigated, with the exception of Columbia-SK group (see Dick, 1953). In South America four viruses have been isolated from mosquitoes, namely, Anopheles A, Anopheles B, Wyeomyia and Ilhéus viruses. Antibody to the latter virus has been found in human beings in the region where this virus was isolated (Laemmert & Hughes, 1947).

All of these viruses, with the exception of Mengo, appear to be distinct entities (Smithburn, 1952; Kerr, 1952). West Nile and Ilhéus viruses fall into the group which includes the St Louis and Japanese B viruses, and there is some evidence that Uganda S virus may be antigenically related to yellow fever (Kerr, 1952). In this respect it is of interest to note that Macnamara (personal communication, 1952) has obtained experimental evidence which suggests that a previous infection with Uganda S virus may modify a subsequent infection with yellow fever.

In addition to these viruses, all of which are believed to infect human beings, strains of Theiler's GD VII virus of mice have been isolated from time to time; one such strain isolated in Uganda and previously referred to as M 1032 virus (Horgan, 1951) was found to be highly pathogenic for cotton-rats when inoculated intracerebrally; Haemagogus A, Haemagogus B, Sabethes and Leucocelaenus isolated in South America (Hughes & Perlowagora, 1950) have all been shown to be strains of the GD VII or FA encephalomyelitis virus of mice (Smithburn, 1952).

References

Balfour, A. (1914) *Lancet*, **1**, 1176
Bates, M. (1944) *Amer. J. trop. Med.* **24**, 91
Bauer, J. H. (1928) *Amer. J. trop. Med.* **8**, 261
Bauer, J. H. & Mahaffy, A. F. (1930a) *Amer. J. Hyg.* **12**, 175
Bauer, J. H. & Mahaffy, A. F. (1930b) *Amer. J. Hyg.* **12**, 155
Beeuwkes, H. & Hayne, T. B. (1931) *Trans. R. Soc. trop. Med. Hyg.* **25**, 107
Beeuwkes, H. & Mahaffy, A. F. (1934) *Trans. R. Soc. trop. Med. Hyg.* **28**, 39
Boshell-Manrique, J. (1948) *Amer. J. trop. Med.* **28**, 457
Bruce-Chwatt, L. J. (1950) *J. Trop. Med. (Hyg.)* **53**, 71
Bugher, J. C. (1951) In: Strode, G. K., ed. *Yellow fever*, p. 299. McGraw-Hill, London
Bugher, J. C., Boshell-Manrique, J., Roca-García, M. & Gilmore, R. M. (1941) *Amer. J. trop. Med.* **21**, 309
Bugher, J. C., Boshell-Manrique, J., Roca-García, M. & Osorno-Mesa, E. (1944) *Amer. J. Hyg.* **39**, 16
Causey, O. R., Kumm, H. W. & Laemmert, H. W., jr (1950) *Amer. J. trop. Med.* **30**, 301
Causey, O. R., Laemmert, H. W., jr & Hayes, G. S. (1948) *Amer. J. Hyg.* **47**, 304
Clark, H. C. (1952) *Amer. J. trop. Med.* **1**, 78
Davis, N. C. (1930a) *Amer. J. Hyg.* **11**, 321
Davis, N. C. (1930b) *J. exp. Med.* **52**, 405
Davis, N. C. & Shannon, R. C. (1929) *J. exp. Med.* **50**, 803
Dick, G. W. A. (1949) *J. Immunol.* **62**, 375
Dick, G. W. A. (1950) *Ann. trop. Med. Parasit.* **44**, 319
Dick, G. W. A. (1952a) *Trans. R. Soc. trop. Med. Hyg.* **46**, 47
Dick, G. W. A. (1952b) *Amer. J. Hyg.* **55**, 140
Dick, G. W. A. (1953) *Trans. R. Soc. trop. Med. Hyg.* **47**, 13
Dick, G. W. A. & Gee, F. L. (1952) *Trans. R. Soc. trop. Med. Hyg.* **46**, 449
Dick, G. W. A. & Haddow, A. J. (1952) *Trans. R. Soc. trop. Med. Hyg.* **46**, 600
Dick, G. W. A. & Horgan, E. S. (1952) *J. Hyg., Camb.* **50**, 376
Dick, G. W. A., Kitchen, S. F. & Haddow, A. J. (1952) *Trans. R. Soc. trop. Med. Hyg.* **46**, 509
Dick, G. W. A., Smithburn, K. C. & Haddow, A. J. (1948) *Brit. J. exp. Path.* **29**, 547
Dunn, L. H. (1927a) *Bull. ent. Res.* **18**, 139
Dunn, L. H. (1927b) *Bull. ent. Res.* **18**, 145
Elton, N. W. (1952a) *Amer. J. publ. Hlth*, **42**, 170
Elton, N. W. (1952b) *Publ. Hlth Rep., Wash.* **67**, 426
Elton, N. W. (1952c) *Amer. J. trop. Med.* **1**, 436
Findlay, G. M. & Broom, J. C. (1933) *Brit. J. exp. Path.* **14**, 391
Findlay, G. M. & MacCallum, F. O. (1937) *Trans. R. Soc. trop. Med. Hyg.* **31**, 103
Findlay, G. M., Stefanopoulo, G. J., Davey, T. H. & Mahaffy, A. F. (1936) *Trans. R. Soc. trop. Med. Hyg.* **29**, 419
Fox, J. P. & Cabral, A. S. (1943) *Amer. J. Hyg.* **37**, 93
Gillett, J. D. (1951) *Ann. trop. Med. Parasit.* **45**, 110
Gillett, J. D., Ross, R. W., Dick, G. W. A., Haddow, A. J. & Hewitt, L. E. (1950) *Ann. trop. Med. Parasit.* **44**, 342
Haddow, A. J. (1945a) *Proc. zool. Soc. Lond.* **115**, 1
Haddow, A. J. (1945b) *Bull. ent. Res.* **36**, 33
Haddow, A. J. (1945c) *Bull. ent. Res.* **36**, 297
Haddow, A. J. (1952) *Ann. trop. Med. Parasit.* **46**, 135
Haddow, A. J. & Dick, G. W. A. (1948) *Ann. trop. Med. Parasit.* **42**, 271
Haddow, A. J., Dick, G. W. A., Lumsden, W. H. R. & Smithburn, K. C. (1951) *Trans. R. Soc. trop. Med. Hyg.* **45**, 189

Haddow, A. J., Gillett, J. D. & Highton, R. B. (1947) *Bull. ent. Res.* **37,** 301
Haddow, A. J. & Mahaffy, A. F. (1949) *Bull. ent. Res.* **40,** 169
Haddow, A. J., Smithburn, K. C., Dick, G. W. A., Kitchen, S. F. & Lumsden, W. H. R. (1948) *Ann. trop. Med. Parasit.* **42,** 218
Hahn, R. G. (1951) *Amer. J. Hyg.* **54,** 50
Horgan, E. S. (1951) *Virus Research Institute annual report, 1950.* Government Printer, Nairobi
Hughes, T. P. (1943) *Trans. R. Soc. trop. Med. Hyg.* **36,** 339
Hughes, T. P. & Perlowagora, A. (1950) *J. Immunol.* **65,** 155
Kerr, J. A. (1932) *Ann. trop. Med. Parasit.* **26,** 119
Kerr, J. A. (1952) *J. Immunol.* **68,** 461
Kirk, R. (1941) *Ann. trop. Med. Parasit.* **35,** 67
Koprowski, H. (1946) *J. Immunol.* **54,** 387
Kumm, H. W. (1950) *Trans. R. Soc. trop. Med. Hyg.* **43,** 673
Kumm, H. W. (1952) *Amer. J. trop. Med.* **1,** 210
Kumm, H. W. & Cerqueira, N. L. (1951) *Bull. ent. Res.* **42,** 195
Kumm, H. W. & Laemmert, H. W., jr (1950) *Amer. J. trop. Med.* **30,** 733
Laemmert, H. W., jr (1946) *Amer. J. trop. Med.* **26,** 33
Laemmert, H. W., jr, Ferreira, L. de C. & Taylor, R. M. (1946) *Amer. J. trop. Med.* **26,** Suppl., p. 23
Laemmert, H. W., jr & Hughes, T. P. (1947) *J. Immunol.* **55,** 61
Lewis, D. J. (1943) *Ann. trop. Med. Parasit.* **37,** 65
Lewis, D. J., Hughes, T. P. & Mahaffy, A. F. (1942) *Ann. trop. Med. Parasit.* **36,** 34
Lloyd, W., Theiler, M. & Ricci, N. I. (1936) *Trans. R. Soc. trop. Med. Hyg.* **29,** 481
Lumsden, W. H. R. (1951) *Bull. ent. Res.* **42,** 317
Macnamara, F. N. (1953) *Trans. R. Soc. trop. Med. Hyg.* **47,** 199
Mahaffy, A. F., Smithburn, K. C. & Hughes, T. P. (1946) *Trans. R. Soc. trop. Med. Hyg.* **40,** 57
Mahaffy, A. F., Smithburn, K. C., Jacobs, H. R. & Gillett, J. D. (1942) *Trans. R. Soc. trop. Med. Hyg.* **36,** 9
Mathis, C., Sellards, A. W. & Laigret, J. (1928) *C. R. Acad. Sci., Paris,* **186,** 604
Melnick, J. L., Paul, J. R., Riordan, J. T., Barnet, V. H., Goldblum, N. & Zabin, E. (1951) *Proc. Soc. exp. Biol., N.Y.* **77,** 661
Peltier, M. (1947) *Amer. J. publ. Hlth,* **37,** 1026
Peltier, M. (1948) In: *Proceedings of the Fourth International Congresses on Tropical Medicine and Malaria, Washington, D.C., May 10–18, 1948,* vol. 1, p. 489. (USA Department of State Publication, No. 3246.) Government Printing Office, Washington
Philip, C. B. (1929) *Amer. J. trop. Med.* **9,** 267
Philip, C. B. (1930a) *Amer. J. trop. Med.* **10,** 1
Philip, C. B. (1930b) *Science,* **71,** 614
Reed, W. & Carroll, J. (1902) *Amer. Med.* **3,** 301
Ross, R. W. & Gillett, J. D. (1950) *Ann. trop. Med. Parasit.* **44,** 351
Sawyer, W. A., Kitchen, S. F., Frobisher, M., jr & Lloyd, W. (1930) *J. exp. Med.* **51,** 493
Sawyer, W. A. & Lloyd, W. (1931) *J. exp. Med.* **54,** 533
Sawyer, W. A. & Whitman, L. (1936) *Trans. R. Soc. trop. Med. Hyg.* **29,** 397
Shannon, R. C., Whitman, L. & Franca, M. (1938) *Science,* **88,** 110
Smithburn, K. C. (1949) *Amer. J. trop. Med.* **29,** 414
Smithburn, K. C. (1952) *J. Immunol.* **68,** 441
Smithburn, K. C. (1953) *Fed. Proc.* **12,** 460
Smithburn, K. C., Goodner, K., Dick, G. W A., Kitchen, S. F. & Ross, R. W. (1949) *Ann. trop. Med. Parasit.* **43,** 182
Smithburn, K. C. & Haddow, A. J. (1944) *J. Immunol.* **49,** 141
Smithburn, K. C. & Haddow, A. J. (1946) *Amer. J. trop. Med.* **26,** 261
Smithburn, K. C. & Haddow, A. J. (1949) *Amer. J. trop. Med.* **29,** 389
Smithburn, K. C. & Haddow, A. J. (1951) *Proc. Soc. exp. Biol., N.Y.* **77,** 130
Smithburn, K. C., Haddow, A. J. & Lumsden, W. H. R. (1949) *Ann. trop. Med. Parasit.* **43,** 74
Smithburn, K. C., Haddow, A. J. & Mahaffy, A. F. (1946) *Amer. J. trop. Med.* **26,** 189
Smithburn, K. C., Hughes, T. P., Burke, A. W. & Paul, J. H. (1940) *Amer. J. trop. Med.* **20,** 471
Smithburn, K. C., Mahaffy, A. F. & Paul, J. H. (1941) *Amer. J. trop. Med.* **21,** 75
Soper, F. L. (1935) *Rev. Hig., Bogota,* **4,** 49
Soper, F. L. (1936a) *Quart. Bull. Hlth Org. L. o. N.* **5,** 19
Soper, F. L. (1936b) *Rev. Hyg., Rio de J.* **10,** 107
Soper, F. L., Penna, H. A., Cardoso, E., Serafim, J., jr, Frobisher, M., jr & Pinheiro, J. (1933) *Amer. J. Hyg.* **18,** 555
South African Institute for Medical Research (1946) *Rep. S. Afr. Inst. med. Res.,* p. 21
Stokes, A., Bauer, J. H. & Hudson, N. P. (1928a) *J. Amer. med. Ass.* **90,** 253
Stokes, A., Bauer, J. H. & Hudson, N. P. (1928b) *Amer. J. trop. Med.* **8,** 103
Taylor, R. M. (1951) In: Strode, G. K., ed. *Yellow fever,* p. 427. McGraw-Hill, London
Taylor, R. M. & Fonseca da Cunha, J. (1946) *Amer. J. trop. Med.* **26,** Suppl., p. 1
Theiler, M. (1930) *Science,* **71,** 367
Theiler, M. (1931) *Ann. trop. Med. Parasit.* **25,** 69
Theiler, M. & Smith, H. H. (1937a) *J. exp. Med.* **65,** 767
Theiler, M. & Smith, H. H. (1937b) *J. exp. Med.* **65,** 787
Wellington, W. G. (1945) *Canad. Ent.* **77,** 7; 21; 44; 69
Whitman, L. (1951) In: Strode, G. K., ed. *Yellow fever,* p. 229. McGraw-Hill, London

* * *

HEPATITIS

F. O. MacCALLUM B.Sc. M.D.

Central Public Health Laboratory
Colindale, London

1 Aetiology
2 Clinical features
3 Pathogenesis
4 Epidemiology
  a Virus A
  b Virus B
5 Immunity
  a Virus A
  b Virus B
6 Control
  a Virus A: general measures of control
  b Virus A: specific prophylaxis
  c Virus B: general measures of control
7 Summary
References

A very brief review of our knowledge of infectious (infective) hepatitis appeared in the first volume of *British Medical Bulletin* in 1943 (MacCallum, 1943). Since that time, a considerable amount of work has been carried out on virus hepatitis, from the standpoint of aetiology, diagnosis, treatment and control, but the advances have unfortunately been much less than one had hoped. The prime factor responsible for this lack of progress has been the failure to find an animal host—other than man—which consistently shows signs, symptoms or lesions pathognomonic of infection with the hepatitis viruses. Various workers (Siede & Meding, 1941; Essen & Lembke, 1944; Henle, Harris, Henle, Harris, Drake, Mangold & Stokes, 1950; Henle, Drake, Henle & Stokes, 1952) have provided evidence that hepatitis viruses have been propagated in the fertile hen-egg, but the proof of this has been obtained only by further inoculation of the egg material into man. Henle and her colleagues (Henle, Drake, Henle & Stokes, 1950) have reported that amniotic fluid infected with at least one of these viruses will, after inactivation by ultra-violet light, give a positive skin reaction when injected intradermally into people who are convalescent from infection with the same virus, but this work has not been confirmed elsewhere. A sufficiently high content of virus or antigen has not yet been obtained in the infective material to give a specific positive reaction with the serum of such convalescents.

Thus the limited advances in knowledge which have been made, as well as the confirmation of various facts deduced by previous epidemiological observations, have been dependent upon the results of experiments on human beings. It must be stressed here that these individuals have been true volunteers of normal mentality who had been informed of the possible risks of morbidity and mortality. In Great Britain the volunteers have been of two types: conscientious objectors in wartime and patients with rheumatoid arthritis. The reasons for using the latter were discussed by MacCallum & Bradley (1944), and will not be mentioned here. After several years' trial it was considered that the results did not justify continuation of the use of the rheumatoid arthritis patients (Gardner, Stewart & MacCallum, 1945). As one can never be certain what previous exposure subjects have had, or of their immune state to these viruses, it is obvious that positive results are easier to interpret than negative ones. A further flaw in many recorded experiments with blood has been the lack of knowledge of the carrier state of the donor and, of course, possible infections from other sources have had to be considered if the recipients were not completely isolated. With these provisos in mind, the author presents his ideas on the conclusions which can be drawn from the various published papers on the subject. In general, these are similar to those found in the Report of the World Health Organization Expert Committee on Hepatitis (1953).[1]

1. Aetiology

As in this World Health Organization (WHO) Report, it is assumed that there are two different viruses, A and B, and thus two types of clinical viral hepatitis (other than that found in yellow fever, Rift Valley fever, and some other suspected virus diseases). The diseases will be described as infectious hepatitis due to virus A and serum (inoculation) hepatitis, which is usually due to virus B, but also may be caused by virus A. The reasons for using these terms may be found in the WHO Report. Although there appear to be at least two different hepatitis viruses, or immunologically distinct strains of the same virus, the clinical disease which they produce in man, and the effect on them of various physical and chemical agents (Table III)—to which they are more resistant than other viruses—are practically the same. It is only in their incubation periods, apparent sites of multiplication, routes of infection, common modes of transmission and immunological behaviour that they differ (Table I). That these forms of hepatitis are due to viruses has been shown by the fact that the disease can

TABLE I.* COMPARISON OF CERTAIN FEATURES OF INFECTIOUS HEPATITIS WITH SERUM HEPATITIS

Feature	Infectious hepatitis virus A	Serum hepatitis virus B
Usual incubation period (days)	15–40	60–160
Virus in faeces	Acute phase	Not demonstrated
Virus in blood	3 days before onset and in acute phase	Incubation period and acute phase
Route of infection (experimental)	Oral and parenteral	Parenteral
Immunity : homologous	Present	Suggested
Immunity : heterologous	None apparent	None apparent
Prophylactic value of γ-globulin	Good	Equivocal or absent
Carrier state†	Unknown	As long as 5 years in blood (one adult)

* Based on Tables I and II of WHO Report (1953).

† Recent report of Capps, Bennett & Stokes (1952) indicated that infants may excrete virus A in faeces for many months during a chronic illness with failure to develop immunity, but that no true healthy carriers have been detected.

[1] Reviewed on p. 242 of this Bulletin.

be produced by bacteria-free filtrates of blood and faeces from patients in the active stage of infectious hepatitis and similar filtrates of blood in serum hepatitis. No organisms have been definitely identified, even under the electron microscope, but one series of filtration experiments has shown that virus B is less than 50 mμ. in its smallest diameter (McCollum, 1952). These viruses have, of course, both been maintained in series experimentally. One of the main points of difference between the two viruses is the fact that virus B has not been recovered from the stools of patients infected with it, but can be maintained only by the parenteral injection of blood. However, no attempt has been made in the incubation period to isolate virus B from the faeces. It may then be a mutant derived from virus A as a result of parenteral passage. There is an analogy in rinderpest, of which one strain in East Africa no longer infects by contact, but only by injection (Hudson, personal communication, 1952). The mechanism of the long incubation period with virus B is quite baffling, for the virus can be isolated from the blood many weeks before the onset of symptoms. Although the long carrier state with virus B may seem unusual, there are other examples of viruses behaving like this in the animal world, e.g., 15 years for virus of equine infectious anaemia of horses and up to 10 years for psittacosis in man and several years in birds; a mosquito infected with yellow fever virus remains infected and unaffected for life. Further discussion of points in Table I are found in other sections. Some of the possibilities of variations in the state of donors of blood and the immune state of the recipients which will govern the results of experiments were not fully appreciated at the time of earlier reports (Cameron, 1943; MacCallum & Bradley, 1944) and the continued quotation of some of these results without this understanding has led to confusion in the minds of those not familiar with all the details of work done to date. In all their experiments in the United States of America with blood from infectious hepatitis, Havens, and Neefe and his colleagues, always obtained short incubation periods, whether they gave it by mouth or parenterally, whereas Cameron in the Middle East obtained incubation periods of one month to six months in six men who received parenteral inoculations of blood from cases of presumed infectious hepatitis in the pyrexial stage. MacCallum and Bradley in England also obtained long incubation periods (64, 75 and 92 days) in one experiment with three adults who received a pool of serum collected from two young soldiers on the day of onset of jaundice in an illness which was considered to be infectious hepatitis. Table II outlines the possible results of experiments with blood in man, depending on hepatitis viruses in inoculum and on the immune state of the recipient, which may explain the above discrepancies. Several persons in the USA and one in Great Britain have been shown by observation, followed by experiment, to be carrying a virus of B type in their serum for several years, with biochemical or histological evidence of liver damage, but without any obvious clinical signs of disease. Persons who have had a history of overt jaundice (presumed virus A but not substantiated) five to ten years or more previously have contributed to pools which have been icterogenic (Cockburn, Harrington, Zeitlin, Morris & Camps, 1951), but that these particular donors were the responsible ones has neither been proved nor disproved. One can only say that attempts to detect a carrier state in the blood after clinical recovery in a very small number of volunteers known to have been infected with virus A have been unsuccessful.

TABLE II. RESULTS WHICH MAY THEORETICALLY BE OBTAINED BY INJECTION OF BLOOD FROM PATIENTS WITH HEPATITIS

O indicates that there is no immunity to hepatitis viruses A or B.

Number of healthy carriers of B unknown, but one estimated figure for one area in England during 1945–48 was not less than 0.35% (Lehane, Kwantes, Upward & Thomson, 1949).

Virus in donor's blood	Immune state of recipient	Result
A	O	Short incubation
A	B	Short incubation
A	A	No disease
B	O	Long incubation
B	A	Long incubation
B	B	No disease
AB*	O	Short + long incubation†
AB	A	Long incubation‡§
AB*	B	Short incubation§
AB*	AB	No disease

* Considered to be A. Carrier state with B not recognizable.
† May be termed a relapse or even a second attack.
‡ Suggested possible example to account for the results of MacCallum & Bradley (1944) which give wrong impressions as to incubation period.
§ Cameron's mixed results (Cameron, 1943) could have been such examples.

2. Clinical Features

Although there may be slight differences between groups of cases of infectious hepatitis and of serum hepatitis, it is impossible to make a differential diagnosis in an individual case. Serum hepatitis can be presumed only if a number of patients have received a common inoculum or received injections at the same session in a clinic, or if one or more transfusions have been given 60–120 days before. It is probable that such patients as the latter will also have received numerous other injections, so that the procedure responsible cannot be determined. The picture of obvious infectious hepatitis with or without jaundice is well known to most physicians and will not be discussed here. The number of persons affected who develop overt jaundice varies from outbreak to outbreak and probably also varies in different age-groups. One of the most striking features of the clinical studies at the present time is the more frequent recognition of occurrence of hepatitis, particularly the more severe form, in babies and young children in whom it has usually been considered to take a benign course. The disease in infants may be either congenitally acquired infection with viruses A or B from the mother's blood or naturally acquired infectious hepatitis contracted at birth or during the neonatal period. Infectious hepatitis may also be acquired in infant nurseries.

There are no data on the approximate size of virus A, but a small number of filtration experiments suggest that virus B is less than 50 mμ. in its smallest diameter (McCollum, 1952), and it is possible that it may pass through the placenta from mother to child. Dible, Hunt, Pugh, Steingold, Marshall & Wood (in press) and Bodian & Newns (1953) have recently described a number of cases of liver disease in

infants in whom the histological changes seen in the liver were indicative of active regeneration following necrosis of the parenchymal cells and were, in general, the same as those seen in infective or serum hepatitis in older patients. As few of the mothers in either series of cases had any history of hepatitis at any time, it was suggested that inapparent serum hepatitis in them may be the most common cause of the infantile disease. Differential diagnosis from congenital obliteration of the bile ducts may be difficult without an exploratory operation. Bodian and Newns have found customary liver function tests in this age-group to be of little value in diagnosis.

Transmission of infectious hepatitis from mother to foetus in the last month of pregnancy has been recorded (Olin, 1947), but Zondek & Bromberg (1947) and Martini (1953) failed to find evidence of hepatitis in new-born infants of 29 and 57 mothers respectively who had jaundice at some time during pregnancy. There is insufficient information at present to determine whether infectious hepatitis in the mother results in congenital deformities, or in an increased rate of still-births, but three of the children born in Martini's series showed brain deformities and about 20% of the survivors in Bodian's group suffered from sequelae of various kinds; two children were deaf and two mentally retarded.

It has been suspected from time to time that young children may also suffer from a milder form of infectious hepatitis, with little or no jaundice, than older children, and in 1945 Pollock demonstrated the occurrence of asymptomatic hepatitis in three children 2–5 years of age in a nursery at the same time that hepatitis with jaundice was present in the young nurses. Recently Capps and his colleagues (Capps, Bennett & Stokes, 1952; Bennett, Capps, Drake, Ettinger, Mills & Stokes, 1952) have described an outbreak of 42 cases of anicteric hepatitis in infants and children under 3 years of age in a nursery where hepatitis with jaundice was endemic among the nurses. Two of the children who had persistent symptoms were found to be excreting virus in their stools 5 and 16 months respectively after onset of symptoms. No evidence of the existence of true healthy carriers has been obtained.

3. Pathogenesis

It has generally been considered that the clinical picture of infectious hepatitis, apart from the jaundice or signs of liver damage, is that found in a systemic infection and, of course, the virus has been recovered from venous blood. It is well recognized that the liver is a very large organ with considerable functional reserve, but is it not possible that too much attention has been focused on the liver as the main site of virus growth? Is it possible that the primary site of virus growth may be in the gastrointestinal tract or the reticulo-endothelial system, and that a toxin is produced which has a particular predilection for the parenchymal cells of the liver? Certainly the microscopic pathological picture of the liver in the acute stage of infective hepatitis is more like that seen in toxic or chemical poisonings than in actual virus infections of the liver such as yellow fever or Rift Valley fever. A picture similar to that found in hepatitis may be seen in glandular fever. Although it is relatively easy to produce toxic changes with chemicals in the livers of experimental animals, no really satisfactory reproducible animal transmissions have been possible with material from patients with infectious hepatitis, but again too much attention may have been focused on the liver. There is little doubt that the reticulo-endothelial and lymphatic systems are affected, at least in some cases, for the spleen may be palpably enlarged; abnormal, large mononuclear cells may be found in smears of the blood; and enlargement of lymphatic glands has not been uncommon in some outbreaks. Whether these changes are the result of actual multiplication of virus in the reticulo-endothelial system is difficult to say.

4. Epidemiology

a. Virus A

The exact mode of spread of this virus is not known, but direct contact with an infected person is probably responsible in most instances and patients appear to be most infectious from a few days before the onset of the symptoms to a few days afterwards.

The results of many of the epidemiological studies in temperate climates in the past led the investigators to conclude that spread was by droplets from the upper respiratory tract in one person to the same site in another. However, attempts to demonstrate the presence of the virus in the pharynx from the onset of symptoms to the onset of jaundice have been unsuccessful and the possibility that the virus is present in or on the pharyngeal mucous membranes before the onset of symptoms, at a time when some patients appear to have been infectious, has not been investigated experimentally.

There is ample experimental proof, as well as epidemiological observation, that the virus is excreted in the faeces during the acute stage of the disease and can infect by the oral route (MacCallum & Bradley, 1944; Havens, Ward, Drill & Paul, 1944; Findlay & Willcox, 1945). Thus fingers, food and fomites can be contaminated with the virus and infection spread either through brief direct contact with the infected individual or indirectly by ingestion of infected food or water or contact with fomites. No evidence of the existence of true healthy carriers has been obtained.

Voegt (1942) and Cameron (1943) and later other investigators in Great Britain and in the USA showed that the virus is present in the blood stream in the acute phase of the disease, so that it is possible for transmission to occur by injection during this period, even in mild unrecognized cases. Blood-sucking insects may act as vectors, but no experimental evidence is available to support this latter suggestion. The epidemiological picture in any community will depend on a combination of the past history of the virus in the community, the present immune, nutritional and metabolic states of the individuals, and the environmental conditions in the home and in the community at large. Thus, in certain countries bordering on the Mediterranean the disease was practically unknown to local physicians, but was obviously endemic, for when the young adults from Germany, Great Britain and North America were introduced, they suffered severely from it in 1942–45. The disease is, of course, also endemic in these latter countries. In a similar fashion, adult male West Africans coming from areas where the virus is known to be present, but among whom the disease is seldom recognized, were quite susceptible when transferred to a new environment. The majority of recognized cases occur in the age-group 5–15 years, although it is known that all age-groups, from birth to death, may be attacked. Observers consider that there has been an unusual number of cases in adults in the past ten years and several outbreaks have been described in which the

rate of attack has been highest in patients over 40 years of age, particularly females, and the mortality rate has been extraordinarily high (20%).

b. Virus B

The meaning of this term, as used in this paper, and the possible nature and origin of the agent, have already been discussed. This virus is apparently transmitted only by parenteral injection. It has been suggested, but not proven, that infection with this virus may occur by the upper respiratory route (Findlay & Martin, 1943; MacCallum, McFarlan, Miles, Pollock & Wilson, 1951) and may be responsible for some sporadic cases of jaundice. It is possible, of course, that this virus may have been maintained for many years only by transplacental transfer and parenteral injection.

The replies to a recent questionnaire sent out to members of the International Society of Geographical Pathology indicated that hepatitis with jaundice had been observed in most parts of the world in patients attending clinics for multiple injections, venepunctures, or superficial bleeding for blood examination. However, only a few reporters admitted having seen cases of jaundice following transfusion, although they agreed it could occur. A possible explanation of this observation, which is in marked contrast to the experience in Britain and in the USA, is that in the latter countries there has been a much greater use of large pools of plasma (from 50–500 donors). This has increased the chances for icterogenic material to be collected and inoculated into a sufficient number of susceptibles for its presence to be detected, and has thus cast suspicion on a common source of their illness. If most recipients are given whole blood only or the plasma from not more than two donors, the chance of obtaining the combination of an infected donor and susceptible recipient is very small. As the donor usually gives no history of jaundice it will be impossible to determine for certain whether the disease in a particular individual has resulted from transfusion, possibly from some injection of anaesthetic or antibiotic, or from contraction of infectious hepatitis by one of the possible modes already described. If a pool consists of plasma from only 8–10 donors, then the probable number of recipients will be only four and the rate may be reduced to little more than when individual bloods are used (Lehane, Kwantes, Upward & Thomson, 1949). British workers have found pools of plasma from 8–10 donors to be satisfactory from the practical standpoint of preparation, but the jaundice rate of 1% to 1.5% is still something to be avoided. Of course, when the rate is low it is unlikely to be recognized in any country unless patients are followed up for four to six months after transfusion.

5. Immunity

a. Virus A

From epidemiological observations there appear to be several strains of virus in different parts of the world which do not completely immunize against each other. However, some of the outbreaks which suggest this may also have been influenced by a change in personal habits in a new environment, as when West Africans went to Italy or to the Far East.

In general, one attack appears to give persistent immunity; only about 3% of patients in several different series of cases gave a history of a second attack of jaundice. Also the serum of convalescents, when given prophylactically, appears to have the power to prevent infection in exposed subjects (Stokes & Neefe, 1945).

b. Virus B

Little is known about the different strains of this virus, but there was some indication that a small group of American soldiers convalescent from yellow fever vaccine jaundice contracted in the USA were immune to an English strain of hepatitis virus B.

The experiments of Neefe, Stokes & Gellis (1945) indicated an homologous immunity and in investigations on subjects without a history of jaundice in England we have never observed an attack rate higher than 75% in any group, and it is usually about 50%. A small number of experiments, however, seem to cast some doubt on the probability of the occurrence of neutralizing antibodies which can be passed to others (Stokes, Blanchard, Neefe, Gellis & Wade, 1948; Henle *et al.* 1952; Henle, personal communication, 1953).

The belief that there are two different viruses, A and B, rests partly on the results of a limited number of experiments, mainly in the USA, in which volunteers were injected with known strains of homologous or heterologous virus after recovery from a previous experimental infection of known type (Havens, 1945; Neefe *et al.* 1945). There has also been a report of observations in the field (Gauld, 1946), in which it was concluded that the attack rate of infectious hepatitis was higher among American troops who had previously had jaundice following yellow fever vaccine (virus B) than among those who had no known previous injection of icterogenic material.

6. Control

a. Virus A: General Measures of Control

The control of an infectious disease in which the aetiological agent cannot be isolated in the laboratory and in which the characteristic clinical feature appears in only a proportion of cases is a very difficult problem. As virtually healthy carriers probably exist (although this is unproven as yet) and patients may be infectious several days before the onset of symptoms, quarantine is of little use. However, a careful watch should be kept for suggestive symptoms such as fever, anorexia and dark urine, and the patient kept at home until recovery has occurred. As the virus has been found in the stools, it is obvious that attention should be paid to personal cleanliness, safe disposal of excreta, and prevention of contamination of food, water and milk supplies, and to possible insect vectors.

b. Virus A: Specific Prophylaxis

There is good evidence that humoral protective antibodies are present in the γ-globulin fraction of the blood of convalescents, and there is evidence that, at least in the USA, γ-globulin prepared from large pools of ordinary adult plasma is capable of terminating infection in an outbreak in a closed community (Stokes & Neefe, 1945; Havens & Paul, 1945). The duration and mechanism of the protection are not clearly defined. Observations made by Capps & Stokes (1952) in a group of nurses inoculated with γ-globulin, and subsequently heavily exposed to hepatitis virus A for a period of several months, suggest that symptomless infection may occur and give protection of a more permanent nature than is usually associated with passive immunization in other virus diseases, such as measles (McDonald & Cockburn, personal communication, 1953).

TABLE III.* COMPARISON OF RESISTANCE OF HEPATITIS VIRUSES A AND B TO PHYSICAL AND CHEMICAL AGENTS

A dash indicates that the test was not done.

Agent	Virus A	Virus B
Temperature resistance		
56° C. (30 min.)	Survived	Survived
60° C. (1 hr.)	—	Survived
60° C. (10 hr.) (albumin)	—	Inactivated
— 10° to — 20° C. (1½ years)	Survived	Survived
— 10° to — 20° C. (4½ years)	—	Survived
Ultra-violet irradiation	—	Equivocal
Chlorine 1 p.p.m. residual (30 min.)	Survived, or inactivated†	—
Tricresol 0.2 %	—	Survived
Phenol-ether (equal parts) 0.5 %	—	Survived
Ether 10 % (24 hr. at 4° C.)	Survived	—
Triple ether extraction of serum	—	Survived
Merthiolate 1 : 2,000	—	Survived
Nitrogen mustard (500 mg./l.)	—	Survived

* From WHO Report (1953) (Table III).
† Inactivation followed adequate coagulation and settling of water.

c. Virus B: General Measures of Control

Here we are dealing with something a little more tangible than in the case of virus A. The main, if not the only, vehicle for transmission of virus B is the blood, which must be injected. Therefore, measures must be directed toward the control of all procedures which bring the blood of one patient into contact with the tissues of another. The two kinds of situation where this occurs are (i) when the blood and blood products are injected for prophylactic or therapeutic purposes and (ii) when instruments are contaminated with blood during the course of blood-taking or injection by intradermal, subcutaneous, intramuscular, or intravenous routes in medical, surgical and dental procedures.

None of the commonly employed antiseptics inactivate hepatitis viruses A or B. In Table III are listed certain data concerning the resistance of both viruses to various chemical and physical agents. Because no mechanism is available for detecting the presence of virus B in a donor's blood the only completely adequate measure for prevention of transmission would be sterilization of the blood. Early work suggested that irradiation with ultra-violet light or treatment with nitrogen mustard could be used for this purpose, but neither has been found to be reliable. Preliminary experiments with plasma containing hepatitis virus B indicate that sulphur mustard (0.005 ml. final concentration) and β-propriolactone (4,000 mg./l.) inactivate this virus, but further experiments will be necessary to show whether such treatment will be of practical value (WHO Report, 1953).

The following recommendations for sterilization of instruments are quoted from the 1953 WHO Report:

> (2) Syringes, needles or other instruments must be thoroughly washed in water immediately after use to prevent organic material coagulating or drying on the surfaces of the instruments and interfering with the effect of subsequent sterilization.
>
> (3) The following methods of sterilization are acceptable: boiling in water for at least 10 minutes, steam under raised pressure (autoclave), and dry-heat sterilization (hot-air oven). If the temperature can be properly controlled in all parts of the hot-air oven, treatment at 170° C. for half an hour is adequate, otherwise treatment at 180° C. for one hour is recommended. Lancets and other instruments used for capillary-blood sampling and scarification may be sterilized in an open flame after washing in cold water.
>
> (4) No chemical disinfectants are accepted for the sterilization of instruments.
>
> (5) When shortage of equipment and staff makes it impossible to resterilize syringes between each injection, for example in mass immunization-campaigns, the needle must be changed and sterilized between each injection, the syringe being resterilized before it is refilled. Although this procedure may reduce the risk of serum hepatitis, it does not eliminate it.

Finally, attention should be drawn to the fact that medical and laboratory personnel handling blood are frequently exposed to the risk of infection by hepatitis viruses and should be instructed in measures to reduce their exposure; Sulkin & Pike (1951) have collected 95 cases of hepatitis in laboratory workers.

7. Summary

A brief review has been made of our present knowledge of virus hepatitis in man and a number of controversial aspects such as aetiology and pathogenesis have been discussed.

Infectious hepatitis is one of the few important virus diseases of man for which no susceptible experimental host has been found. It is a serious cause of morbidity, particularly in times of stress, in populations with a high standard of hygiene, in much the same way as poliomyelitis, and is worthy of more intensive efforts to isolate the virus for study and development of prophylactic measures.

Certain hepatitis viruses may circulate in the blood stream for long periods of time without causing obvious illness and may easily be transferred to other persons who are susceptible and develop the disease. For this reason it is essential that transfusions of blood should be given only when essential, and extreme care must be taken to sterilize adequately all instruments used for injections or collection of blood. Meanwhile, efforts must be intensified to find methods of detecting the healthy carriers and inactivating the viruses in blood and blood products without injuring the blood.

γ-Globulin from pools of adult plasma carries practically no risk of transmitting hepatitis and may be used prophylactically against infectious hepatitis. This same γ-globulin, and even that from convalescents who are known to have had serum hepatitis, is not a satisfactory prophylactic against serum hepatitis.

References

Bennett, A. M., Capps, R. B., Drake, M., Ettinger, R. H., Mills, E. H. & Stokes, J. (1952) *Arch. intern. Med.* **90,** 37
Bodian, M. & Newns, G. H. (1953) *Arch. franç. Pédiat.* **10,** 169
Cameron, J. D. S. (1943) *Quart. J. Med.* n.s. **12,** 139
Capps, R. B., Bennett, A. M. & Stokes, J. (1952) *Arch. intern. Med.* **89,** 6
Capps, R. B. & Stokes, J. (1952) *J. Amer. med. Ass.* **149,** 557
Cockburn, W. C., Harrington, J. A., Zeitlin, R. A., Morris, D. & Camps, F. E. (1951) *Brit. Med. J.* **2,** 6
Dible, J. H., Hunt, W. E., Pugh, V., Steingold, L., Marshall, A. G. & Wood, J. H. F. (In press) *J. Path. Bact.*
Essen, K. W. & Lembke, A. (1944) *Med. Z.* **1,** 99
Findlay, G. M. & Martin, N. H. (1943) *Lancet,* **1,** 678
Findlay, G. M. & Willcox, R. R. (1945) *Lancet,* **1,** 212

Continued at foot of page 226

THE PSITTACOSIS-LYMPHOGRANULOMA GROUP OF VIRUSES

S. P. BEDSON M.D. D.Sc. F.R.C.P. F.R.S.

Bland-Sutton Institute of Pathology
Middlesex Hospital, London

1 The psittacosis-lymphogranuloma group in general
2 Psittacosis
  a Natural history of psittacosis
  b Diagnosis of psittacosis in birds and in man
  c Control of psittacosis
  d Treatment of psittacosis
3 Lymphogranuloma venereum
  a The virus
  b Natural history of lymphogranuloma venereum
  c Laboratory diagnosis of lymphogranuloma venereum
  d Treatment of lymphogranuloma venereum
4 Trachoma and inclusion conjunctivitis
References

1. The Psittacosis-Lymphogranuloma Group in General

The work of the last twelve years has been responsible for adding considerably to the list of viruses belonging to the psittacosis-lymphogranuloma group.

In 1939 there were four: psittacosis, lymphogranuloma venereum, and the two somewhat less closely related viruses of trachoma and inclusion conjunctivitis. Today we have, in addition, strains of psittacosis virus differing sufficiently from the classical parrot strain to merit, in the judgement of some workers, the distinctive name of ornithosis, and a number of mammalian viruses. The latter comprise a virus of mouse pneumonitis (Nigg, 1942; Gönnert, 1942), the viruses of cat pneumonia (Baker, 1944), and of enzootic abortion of ewes (Stamp, McEwen, Watt & Nisbet, 1950), a virus found in the intestinal tract of calves (York & Baker, 1951) and one isolated from diseased opossums in Colombia (Roca-García, 1949). One ought perhaps also to mention certain strains of psittacosis virus isolated from outbreaks of atypical pneumonia in man—the San Francisco, Louisiana and Illinois strains—the avian origin of which was never established. The virulence of these particular strains and the readiness with which they spread from man to man has led to the suggestion that they might be man-adapted strains of psittacosis virus. In this connexion it should, however, be mentioned that Meyer & Eddie (1952) drew attention to the similarity between the virus obtained by Olson & Larson (1944) from patients in the pneumonitis outbreak of the Bayou country of Louisiana and a strain of psittacosis virus isolated from snowy egrets in the same part of the USA. The suggestion is, of course, that the Louisiana outbreak studied by Olson and Larson did, in fact, originate in birds. All these viruses have characters in common. They are similar in size and when multiplying they pass through a developmental cycle similar to that first described in the case of psittacosis virus (Bedson & Bland, 1932, 1934). They are basophilic and, like the rickettsiae, they stain readily by the methods of Castaneda and Macchiavello and, with the possible exception of the viruses of trachoma and inclusion blenorrhoea, they share an antigen in common, a heat-stable important constituent which tends to dominate serological reactions with these viruses. Most authorities today are of the opinion that the agents of this group are more closely related to the rickettsiae than they are to the viruses. The affinity for basic dyes shared by rickettsiae and the viruses of the psittacosis-lymphogranuloma group distinguishes them from what are sometimes referred to now as the " true " viruses and suggests some common chemical constituent, a deoxyribonucleic acid according to Lépine & Sautter (1946). And the available evidence suggests that the psittacosis-lymphogranuloma agents, like the rickettsiae, multiply by binary fission, whilst few today think that the true viruses make use of this procedure. Recognition of this relationship is responsible for the proposal of Moshkovsky (1945) that the agents of the psittacosis-lymphogranuloma group should form one of the families (Chlamydozoaceae) of the order Rickettsiales. Before we leave the question of the composition of this group—the Chlamydozoaceae—some reference should be made to the agent of cat-scratch fever. This condition, first recognized and described in France in recent times (Mollaret, Reilly, Bastin & Tournier, 1950) is now known to occur in other parts of the world, including

---

HEPATITIS

Continued from page 225

Gardner, F., Stewart, A. & MacCallum, F. O. (1945) *Brit. med. J.* **2,** 677
Gauld, R. (1946) *Amer. J. Hyg.* **43,** 310
Havens, W. P., jr (1945) *Proc. Soc. exp. Biol., N.Y.* **59,** 148
Havens, W. P., jr & Paul, J. R. (1945) *J. Amer. med. Ass.* **129,** 270
Havens, W. P., jr, Ward, R., Drill, V. A. & Paul, J. R. (1944) *Proc. Soc. exp. Biol., N.Y.* **57,** 206
Henle, G., Drake, M. E., Henle, W. & Stokes, J. (1950) *Proc. Soc. exp. Biol., N.Y.* **73,** 603
Henle, W., Drake, M. E., Henle, G. & Stokes, J. (1952) *Arch. ges. Virusforsch.* **4,** 612
Henle, W., Harris, S., Henle, G., Harris, T. A., Drake, M. E., Mangold, F. & Stokes, J. (1950) *J. exp. Med.* **92,** 271
Lehane, D., Kwantes, C. M. S., Upward, M. G. & Thomson, D. R. (1949) *Brit. med. J.* **2,** 572
MacCallum, F. O. (1943) *Brit. med. Bull.* **1,** 112
MacCallum, F. O. & Bradley, W. H. (1944) *Lancet,* **2,** 228
MacCallum, F. O., McFarlan, A. M., Miles, J. A. R., Pollock, M. R. & Wilson, C. (1951) *Spec. Rep. Ser. med. Res. Coun., Lond.* No. 273
McCollum, R. (1952) *Proc. Soc. exp. Biol., N.Y.* **81,** 157
Martini, G. A. (1953) *Schweiz. Z. Path.* **16,** 475
Neefe, J. R., Stokes, J. & Gellis, S. S. (1945) *Amer. J. med. Sci.* **210,** 561
Olin, G. (1947) *Acta med. scand.* Suppl. No. 196, p. 381
Pollock, M. R. (1945) *Brit. med. J.* **2,** 598
Siede, W. & Meding, G. (1941) *Klin. Wschr.* **20,** 1065
Stokes, J., Blanchard, M., Neefe, J. R., Gellis, S. S. & Wade, G. R. (1948) *J. Amer. med. Ass.* **138,** 336
Stokes, J. & Neefe, J. R. (1945) *J. Amer. med. Ass.* **127,** 144
Sulkin, S. E. & Pike, R. M. (1951) *Amer. J. publ. Hlth,* **41,** 769
Voegt, H. (1942) *Münch. med. Wschr.* **89,** 76
World Health Organization. Expert Committee on Hepatitis (1953) *World Hlth Org. techn. Rep. Ser.* No. 62
Zondek, B. & Bromberg, Y. M. (1947 *J. Mt Sinai Hosp.* **14,** 222

Britain. It seems clearly to be an infective process which is initiated frequently, though by no means always, by the scratch of a cat; some cases follow scratches by thorns. After an interval of a week or so there is enlargement of the lymph glands draining the cutaneous area involved, associated with some general symptoms and a low-grade fever. Commonly it is the epitrochlear lymph nodes which are affected and, after enlarging considerably, they become fluctuant and break down. The pus is culturally sterile and contains no demonstrable microbe; with the possible exception of the production of cutaneous lesions in the skin of monkeys and of one human volunteer by the intradermal injection of the pus (Mollaret, Reilly, Bastin & Tournier, 1951), transmission experiments have been negative. With the heated pus, however, it is possible to elicit a positive skin reaction of the delayed or tuberculin type in a high percentage of cases (Mollaret *et al.* 1950). The histological changes produced in the affected lymph nodes resemble closely those found in lymphogranuloma venereum, and Mollaret and his colleagues claim to have demonstrated virus particles similar to those of the psittacosis-lymphogranuloma group in sections of the human lymph glands and of the cutaneous lesions produced in monkeys. This, together with the demonstration that the serum from a number of these cases has given a positive complement fixation with a psittacosis-lymphogranuloma group antigen, has suggested that the agent of cat-scratch fever belongs to this group. In this connexion it should be noted that the Frei test is negative in cat-scratch fever and that the heated cat-scratch pus has failed to evoke a reaction in those cases of lymphogranuloma venereum which have been so tested. Clearly this postulated relationship of the agent of cat-scratch fever is far from being established.

2. Psittacosis

Contrary to the belief held in 1930, when the virus aetiology of psittacosis was established, it is not just a single virus which is concerned but a number of closely related viruses. In addition to the strains of psittacine origin, there are those found naturally infecting other species of birds, to which the name ornithosis is commonly given and which, generally speaking, are of less virulence for man. And as already mentioned, there are the strains isolated from human outbreaks apparently unconnected with birds, the San Francisco, Illinois and Louisiana viruses, which, it has been suggested, are possibly strains of psittacosis virus that have become adapted to man and use him as their principal host. The question also arises as to whether or not the viruses of feline and mouse pneumonitis should be included in the psittacosis group. Though obviously closely related to the psittacosis viruses, as indeed are also the virus isolated from the intestinal tract of calves by York & Baker (1951), the opossum virus of Roca-García (1949) and the virus of enzootic abortion of ewes (Stamp *et al.* 1950), it seems most probable that these mammalian viruses represent separate species.

a. Natural History of Psittacosis

Psittacosis is primarily a disease of birds and, although man is susceptible, his infection is, relatively speaking, a rare event and of no consequence as far as the continued survival of the virus is concerned. We know today that the avian hosts of these viruses extend far beyond the birds of the parrot family; in fact, as Meyer (1948) states, at least 70 species of birds belonging to ten orders, including the psittaciformes, have been found naturally to harbour the infection. These include pigeons and doves, the domestic fowl, duck and turkey, the fulmar petrel, sea-gulls and various species of finch, many of which have, at times, been responsible for cases of human psittacosis. The young bird acquires infection from its parents while it is still in the nest and commonly this infection is symptomless; a carrier state is set up which can persist for years or even for the rest of the life of the bird. Sometimes the naturally occurring infection may lead to devastating epizootics. Burnet (1935, 1939) has described such an event among parrots and parakeets in Australia and more recently Parodi & Silvetti (1946) witnessed a similar happening in the Argentine. But these are rare events and the naturally acquired infection remains symptomless unless something disturbs the balance between host and parasite and allows the latter to gain the upper hand. This is well illustrated by the fate which frequently befalls newly caught parrots or parakeets when they are sent on long journeys under unsatisfactory conditions. Despatched in apparently good health, many of the birds may be diseased and some even dead or dying when they reach their destination. The latent infection in some of the birds is lit into activity and rapidly spreads to the uninfected birds in the consignment. Another excellent example of the activation of latent psittacosis infection was provided by the chance observation of Pinkerton & Swank (1940) that some of their experimental pigeons fed on a diet deficient in thiamine developed psittacosis. It was not known at the time that many pigeon stocks carried a latent infection with a psittacosis virus; in fact this observation of Pinkerton and Swank was the first intimation that pigeons are one of the natural hosts of this virus. The vast majority of human infections with the psittacosis viruses are derived directly from birds which are either carrying the infection or are manifestly diseased. The birds excrete the virus in the droppings and these when dried provide an infective dust which is inhaled by man, and since the acutely ill bird not only excretes much more virus but also suffers from diarrhoea, as a source of human infection it is much more to be feared than the apparently healthy carrier. It is for this reason that practically all serious outbreaks of psittacosis in man have had their origin in recently imported birds of the parrot family.

b. Diagnosis of Psittacosis in Birds and in Man

In both the bird and in man there is nothing sufficiently distinctive about the clinical picture of psittacosis to make it readily recognizable and for accurate diagnosis the help of the laboratory is essential. The laboratory methods available for this purpose have undergone little change in recent years. In suspected psittacosis of birds, isolation of the virus from their tissues and more particularly from the spleen is the main objective. This, of course, entails the sacrifice of the birds; this, if they were ill, would be necessary in any case. Where, however, it is a question of detecting latent infection in a bird stock, such as a flock of pigeons or a consignment of parrots or parakeets, and it is not possible to sacrifice some of the birds for virus investigation, the complement-fixation test may prove useful (Meyer & Eddie, 1939). Although the correlation between presence of antibody in the serum and of virus in the tissues is not as close as was at one time believed (Davis, 1950), the demonstration of

antibodies in pooled serum samples is valuable evidence of the presence of latent psittacosis infection in a flock of birds. In this connexion it should perhaps be mentioned that, although pigeon and parrot sera work satisfactorily with the ordinary or direct complement-fixation test, this is not so in the case of the serum of the domestic fowl, duck or turkey; for them the indirect complement-fixation test must be used (Rice, 1947; Karrer, Meyer & Eddie, 1950). Although the aim of the laboratory in suspected psittacosis in man, just as in the case of the bird, is the isolation of the virus, the opportunities for attempting this occur less frequently. The virus is present in the blood in the early days of the disease but it is unusual for psittacosis to be suspected at this stage. Later in the disease, in those cases that develop pneumonic foci and produce sputum, the virus may be looked for in this material. But it is by no means generally recognized that in a considerable percentage of cases of human psittacosis the lungs are not involved and even in those with pneumonic foci sputum may be extremely scanty or even absent. So that in practice one has to rely largely on serological investigation—the complement-fixation test—in the diagnosis of human infections with the psittacosis viruses. Knowledge of the antigenic structure of these viruses (Bedson, Barwell, King & Bishop, 1949; Barwell, 1952) has made it clear that it is the heat-stable antigen common to the viruses of the psittacosis-lymphogranuloma group which is the operative one in the complement-fixation test as practised since its introduction 18 years ago. Although specific antibodies are, of course, produced in these human infections and their detection by the complement-fixation test is possible (Bedson *et al.* 1949), technical difficulties make this test unsuitable for routine use (Monsur & Barwell, 1951). Therefore it is the group complement-fixation test which is employed today, and provided that its limitations are kept in mind it gives valuable service.

c. Control of Psittacosis

It cannot be claimed that psittacosis in pigeons, domestic fowls, ducks or turkeys on farms or breeding establishments in Great Britain presents a problem of major importance either from the economic or public health point of view. Should the infection, which is usually inapparent, get out of hand and assume epidemic form, this almost invariably means some departure from good husbandry, and rectifying this, together with destruction of the diseased birds, will usually bring the outbreak within control. Psittacosis in birds of the parrot family is more important because, as pets, they come into close relation with man, and also because the strains of virus found in them seem to be more virulent than those from other kinds of birds. Therefore, something is to be said for attempting the control of the disease in psittacine birds and with the knowledge at our disposal it should be possible to clean up the breeding stocks of these birds. Legislation prohibiting the importation of parrots and parakeets, except under licence and with an adequate period of quarantine, would then prevent the reintroduction of infection. There is no doubt that the serious outbreaks of psittacosis in man have had their origin in recently imported parrots and parakeets and this, as we have seen, is due to the fact that these birds often reach their destination in a heavily diseased state. The embargo placed on the importation of these birds by Britain and the USA and other countries after the pandemic of 1930 has undoubtedly reduced materially the incidence of human psittacosis and, although we now know that the disease is enzootic in bird stocks in these countries, it would still be a mistake completely to free the traffic in psittacine birds. We have indeed had proof of this. Britain raised the embargo in January 1952 and within a few months several outbreaks of psittacosis originating in imported parrots had occurred. The embargo has had to be reimposed. Drachman (1953) reports that, as the result of lifting the embargo in New York State in January 1952, Upper Westchester County, which had been without a case of human psittacosis for ten years, experienced two outbreaks originating in imported psittacine birds.

d. Treatment of Psittacosis

We are no longer confined to the symptomatic treatment of human psittacosis, as these viruses are susceptible to a number of the antibiotics—penicillin, chloromycetin, aureomycin and terramycin. A few of the strains are even sulphonamide-sensitive, too few, however, to make treatment with sulphonamides worth considering, though the literature contains reports of cases of psittacosis successfully treated with these drugs. Penicillin has been used with success, but the studies of Hurst, Peters & Melvin (1950) on the chemotherapy of experimental psittacosis[1] have clearly indicated the superiority of aureomycin and terramycin; chloramphenicol proved the least effective. Experimental work in mice (Bedson & May, 1945; Hurst, 1948) has shown that treatment with penicillin, even when started at the time of infection, although suppressing all clinical signs of disease, does not eradicate the infection. The work of Gogolak & Weiss (1950) has also shown that the effect of aureomycin on the nearly related viruses of feline and mouse pneumonitis is viristatic but not viricidal; it prevents division of the initial bodies. These findings explain the failure of these drugs to clear up the carrier state in birds infected with psittacosis (Quan, Meyer & Eddie, 1950); they also suggest that to be effective therapeutically they should be given in as large a dose and as early in the disease as possible.

3. Lymphogranuloma Venereum

a. The Virus

Although the filtrable nature of the agent of this disease was established in 1930 by the work of the Swedish investigators Hellerström & Wassén, there was a lapse of five years before it was realized that this virus when stained could be seen with the ordinary microscope (Miyagawa, Mitamura, Yaoi, Ishii, Nakajima, Okanishi, Watanabe & Sato, 1935). Confirming this, Findlay, Mackenzie & MacCallum (1938) showed also that the virus when multiplying passed through a developmental cycle essentially like that of psittacosis virus. The estimate of the size of this virus, made by means of gradocol filters (Broom & Findlay, 1936), differs considerably from the measurements made on electron micrographs (Kurotchkin, Libby, Gagnon & Cox, 1947)—125–175 mμ. as against 430 ± 47 mμ.—but a similar discrepancy has been found with other viruses. Reference has already been made to the heat-stable antigens possessed by this virus in common with the other agents of the psittacosis-lymphogranuloma group and to the fact that this group antigen is responsible for the very high level of cross-reactions given by these viruses and their antisera in the complement-fixation test. The possession of specific antigens by lymphogranuloma venereum virus and

[1] See Hurst, p. 180 of this Bulletin.

other members of the group has been shown by means of the neutralization test (Hilleman, 1945) and by the complement-fixation reaction (Bedson *et al.* 1949; Monsur & Barwell, 1951). Rake & Jones (1944) have also shown that this virus, like others of the group, when grown in the yolk sac is highly toxic for the mouse, particularly when administered intravenously. This toxin is very labile, it is intimately associated with the virus particle—it is an endotoxin—and it is specific. Antisera prepared in the domestic fowl or rabbit with formolized virus neutralize specifically these toxins produced by the viruses of the psittacosis-lymphogranuloma group and this fact can be made use of in the identification of the species (Manire & Meyer, 1950). It seems probable that the endotoxins and the labile specific antigen demonstrable by the neutralization and complement-fixation tests are one and the same.

b. Natural History of Lymphogranuloma Venereum

This disease, which is of world-wide distribution and to which all races appear to be susceptible, is peculiar to the human species. Certain species of animals—monkeys, white mice, guinea-pigs and chick embryos—can be infected experimentally, but it is in man alone that natural infection occurs. It is spread by venereal contact, though probably not exclusively so, since it is known that this virus can attack the conjunctiva and that an oculoglandular syndrome occurs in lymphogranuloma venereum, which suggests that infection may at times be transmitted innocently. The similarity presented by the viruses of the psittacosis-lymphogranuloma group in size, staining properties, developmental cycle and antigenic constitution suggests that they have arisen from a common ancestral stock; one is tempted to believe that early in their history they existed as parasites of certain bird species and that it is from the avian source that the mammalian species, including the three which have man as their sole host —the virus of lymphogranuloma venereum, trachoma and inclusion blenorrhoea—have originated. If this is so, it would seem that the last two became adapted to man earlier in the history of the group than the virus of lymphogranuloma venereum, since they are more restricted in their experimental host range and the disturbance which they produce in man when they infect is minimal.

c. Laboratory Diagnosis of Lymphogranuloma Venereum

Just as in psittacosis, the group complement-fixation test is widely used now in the diagnosis of lymphogranuloma venereum; there is no doubt that it is a valuable aid (Bedson *et al.* 1949). The test for skin hypersensitivity introduced by Frei 18 years ago is also relied on extensively in the investigation of suspected cases of lymphogranuloma venereum. It is still not fully appreciated that this test, as devised by Frei and practised today, is only group-specific (Eaton, Beck & Pearson, 1941; Bedson *et al.* 1949). As Barwell (1952) has shown, however, treatment of the virus with dilute acid—0.02 N-hydrochloric acid—extracts an antigenic fraction which, though not reactive in serological tests in vitro, none the less elicits a reaction in the skin of patients when it is injected intradermally. And a limited trial of such extracts made from the viruses of lymphogranuloma venereum and psittacosis has shown that, unlike the Frei antigen, the reactions they evoke are either strictly specific or very largely so. This specific antigenic fraction appears stable and will no doubt replace the Frei antigen when it has been tried out more extensively.

d. Treatment of Lymphogranuloma Venereum

Lymphogranuloma venereum virus was one of the first of this family shown to be susceptible to sulphonamides; in fact, in the opinion of some venereologists, treatment with the sulphonamides is still the method of choice in lymphogranuloma venereum. This is not due to any lack of susceptibility of the virus to the antibiotics; the contrary has been demonstrated by Hurst *et al.* (1950), and aureomycin has been used with good results in this disease. The reason for this conservative attitude is the fear of masking an infection with *Treponema pallidum*, since the dose of penicillin or aureomycin which would be adequate in the treatment of lymphogranuloma venereum might only suppress the syphilitic infection; this risk is avoided if sulphonamides are used, since they are without effect on the treponema. It would seem, however, that aureomycin is the treatment of choice in lymphogranuloma venereum.

4. Trachoma and Inclusion Conjunctivitis

These two viruses were the first of the family to be seen, though at the time that this was recorded over 40 years ago few believed that the particles stained and seen were, in fact, the agents of these diseases. This has since been established. The viruses of trachoma and of inclusion conjunctivitis look like and stain like other members of this family and they both exhibit the same sequence of developmental forms when they multiply (Thygeson, 1934a, 1934b). It is also claimed (Rake, Shaffer & Thygeson, 1942) that they possess the same group antigen as the rest of this group of viruses, since sera from cases of trachoma and inclusion conjunctivitis will fix complement with heated antigens made from psittacosis and lymphogranuloma venereum viruses. It should be noted, however, that the titres obtained were never high and that the Frei test is negative in trachoma, so it is perhaps premature to accept as established that they share the heat-stable antigen which is such a striking common feature of the psittacosis-lymphogranuloma viruses. This, and a number of other questions relating to the viruses of trachoma and inclusion conjunctivitis, will find no ready solution as long as the only alternatives to man as an experimental animal for the study of these viruses are certain species of apes. This has been a heavy handicap, particularly so since it has also not been found possible to grow either virus in tissue culture or in the developing egg. It is true that the recent literature contains claims to the contrary. Macchiavello (1944) states unequivocally that he has succeeded in establishing trachoma virus in the yolk sac of the developing hen-egg, and Poleff (1949) confirms this; Stewart & Badir (1950), alternating inoculation of the egg with infection of the monkey, have also reported several successful passages of this virus. Moreover, Arakawa & Kitamura (1950) claim to have isolated eight strains of trachoma virus by inoculating mice intracerebrally with trachomatous material.

One wonders if these claims will prove substantial. In this connexion it is perhaps significant that we still await the answers to many questions concerning these two viruses and more particularly concerning their relationship to the psittacosis-lymphogranuloma viruses and even to one another.

References

Arakawa, S. & Kitamura, O. (1950) *Jap. J. exp. Med.* **20,** 579
Baker, J. A. (1944) *J. exp. Med.* **79,** 159
Barwell, C. F. (1952) *Brit. J. exp. Path.* **33,** 268
Bedson, S. P., Barwell, C. F., King, E. J. & Bishop, L. W. J. (1949) *J. clin. Path.* **2,** 241
Bedson, S. P. & Bland, J. O. W. (1932) *Brit. J. exp. Path.* **13,** 461
Bedson, S. P. & Bland, J. O. W. (1934) *Brit. J. exp. Path.* **15,** 243
Bedson, S. P. & May, H. B. (1945) *Lancet*, **2,** 394
Broom, J. C. & Findlay, G. M. (1936) *Brit. J. exp. Path.* **17,** 135
Burnet, F. M. (1935) *J. Hyg., Camb.* **35,** 412
Burnet, F. M. (1939) *Med. J. Aust.* **1,** 545
Davis, D. J. (1950) *J. Amer. vet. med. Ass.* **116,** 220
Drachman, T. S. (1953) *Amer. J. publ. Hlth*, **43,** 165
Eaton, M. D., Beck, M. D. & Pearson, H. E. (1941) *J. exp. Med.* **73,** 641
Findlay, G. M., Mackenzie, R. D. & MacCallum, F. O. (1938) *Trans. R. Soc. trop. Med. Hyg.* **32,** 183
Gönnert, R. (1942) *Zbl. Bakt.* (1 Abt., Orig.) **148,** 331
Gogolak, F. M. & Weiss, E. (1950) *J. infect. Dis.* **87,** 264
Hellerström, S. & Wassén, E. (1930) *Int. derm. Congr.* **8,** 1147
Hilleman, M. R. (1945) *J. infect. Dis.* **76,** 96
Hurst, E. W. (1948) *Brit. J. Pharmacol.* **3,** 18
Hurst, E. W., Peters, J. M. & Melvin, P. (1950) *Brit. J. Pharmacol.* **5,** 611
Karrer, H., Meyer, K. F. & Eddie, B. (1950) *J. infect. Dis.* **87,** 13; 24
Kurotchkin, T. J., Libby, R. L., Gagnon, E. & Cox, H. R. (1947) *J. Immunol.* **55,** 283
Lépine, P. & Sautter, V. (1946) *Ann. Inst. Pasteur*, **72,** 174
Macchiavello, A. (1944) *Rev. ecuator Hig.* **2,** 211
Manire, G. P. & Meyer, K. F. (1950) *J. infect. Dis.* **86,** 241
Meyer, K. F. (1948) In : Biester, H. E. & Schwarte, L. H., ed. *Diseases of poultry*, 2nd ed., p. 513. Iowa State College Press, Ames, Iowa
Meyer, K. F. & Eddie, B. (1939) *J. infect. Dis.* **65,** 225
Meyer, K. F. & Eddie, B. (1952) *Acta trop., Basel*, **9,** 204
Miyagawa, Y., Mitamura, T., Yaoi, H., Ishii, N., Nakajima, H., Okanishi, J., Watanabe, S. & Sato, K. (1935) *Jap. J. exp. Med.* **13,** 1
Mollaret, P., Reilly, J., Bastin, R. & Tournier, P. (1950) *Pr. méd.* **58,** 1535
Mollaret, P., Reilly, J., Bastin, R. & Tournier, P. (1951) *Pr. méd.* **59,** 701
Monsur, K. A. & Barwell, C. F. (1951) *Brit. J. exp. Path.* **32,** 414
Moshkovsky, S. D. (1945) *Usp. Sovr. Biol.* **19,** 1
Nigg, C. (1942) *Science*, **95,** 49
Olson, B. J. & Larson, C. L. (1944) *Publ. Hlth Rep., Wash.* **59,** 1373
Parodi, A. S. & Silvetti, L. M. (1946) *Prensa méd. argent.* **33,** 529
Pinkerton, H. & Swank, R. L. (1940) *Proc. Soc. exp. Biol., N.Y.* **45,** 704
Poleff, L. (1949) *Ann. Inst. Pasteur*, **77,** 322
Quan, S. F., Meyer, K. F. & Eddie, B. (1950) *J. infect. Dis.* **86,** 132
Rake, G. & Jones, H. P. (1944) *J. exp. Med.* **79,** 463
Rake, G., Shaffer, M. F. & Thygeson, P. (1942) *Proc. Soc. exp. Biol., N.Y.* **49,** 545
Rice, C. E. (1947) *Canad. J. comp. Med.* **11,** 236
Roca-García, M. (1949) *J. infect. Dis.* **85,** 275
Stamp, J. T., McEwen, A. D., Watt, J. A. A. & Nisbet, D. I. (1950) *Vet. Rec.* **62,** 251
Stewart, F. H. & Badir, G. (1950) *J. Path. Bact.* **62,** 457
Thygeson, P. (1934a) *Amer. J. Ophthal.* **17,** 1019
Thygeson, P. (1934b) *Arch. Ophthal., N.Y.* **12,** 307
York, C. J. & Baker, J. A. (1951) *J. exp. Med.* **93,** 587

* * *

Q FEVER IN BRITAIN

M. G. P. STOKER M.D.

Huddersfield Lecturer in Special Pathology
University of Cambridge

1 The causative organism
2 Clinical aspects
3 Prevalence of Q fever in Great Britain
4 Sources and mode of infection
5 Conclusion
References

For many years Great Britain has been considered free of the rickettsial diseases of man except for occasional imported cases of murine or epidemic typhus (including Brill's disease). Recently, however, Q fever has been recognized as a widespread rickettsial infection of man and domestic animals and the cause of a febrile illness in the former often resembling primary atypical pneumonia.

Q fever was originally discovered in Brisbane abattoir workers by Derrick (1937) and it is now known to be widely distributed throughout the world in a great variety of animal and arthropod hosts (see reviews by Parker, Bell & Stoenner, 1949; Marmion, 1951). This paper, however, is principally concerned with work carried out in Great Britain and it does not attempt to cover the intensive investigations in other parts of the world such as Australia (Derrick, 1944), the Mediterranean region (Caminopetros, 1948; Robbins & Ragan, 1946; Blanc, Martin, Bruneau & Maurice, 1948), and California (Huebner & Bell, 1951; Lennette & Clark, 1951).

First evidence of Q fever in Britain was produced when Q fever antibodies were found in sera from 3 out of 24 patients with atypical pneumonia (Stoker, 1949). Soon afterwards MacCallum, Marmion & Stoker (1949) isolated a strain of the causative organism, *Rickettsia burneti*, from one of the patients in an outbreak of Q fever at the Royal Cancer Hospital, London (Harman, 1949).

The principal source of infection in this hospital outbreak was a fatal case of Q fever admitted from a town in north-east Kent (Marmion & Stoker, 1950) and attention was drawn to this county, where Q fever was soon found to be endemic. Two other outbreaks were detected, one involving 28 members of the Canterbury School of Art (Harvey, Forbes & Marmion, 1951) and the other 16 persons at a farm in south Kent (Stoker & Thompson, 1953). Marmion, Stoker, McCoy, Malloch & Moore (1953) have also described 69 sporadic cases of Q fever occurring in Great Britain, and although nearly half were infected in Kent, the remainder were scattered over a wide area of England and Wales. It is not known how much the numbers in Kent are influenced by the intensive investigation which followed chance identification of the first cases in that county, but it is of interest that no clinical Q fever has yet been detected in East Anglia, despite a search starting from Cambridge.

Many of the investigations described here have been carried out with valuable help from members of the Public Health Laboratory Service and from veterinary surgeons of the Ministry of Agriculture.

1. The Causative Organism

Although *R. burneti* resembles other pathogenic rickettsiae in its morphology, staining reactions and growth requirements, it differs from them in its ability to withstand adverse conditions of temperature and humidity and in its resistance to chemical agents. These properties probably account for the ability of the organism to maintain itself in nature without an arthropod vector and make possible the air-borne dissemination of infection. *R. burneti* will also pass filters which will hold back other rickettsiae, but this may also be due in part to continued viability during the conditions of filtration, as there is no great difference in size judged by electron microscopy.

The resistance of *R. burneti* to the conditions used in pasteurization is of obvious importance because of the presence of the organism in raw milk. Marmion, Rowlands, MacCallum & Thiel (1951) investigated quantitatively the effect of the High-Temperature: Short-Time and vat-holder techniques upon naturally and artificially infected milk. The results show that the standard temperatures used in Great Britain are only at the border-line of safety but that the low infectivity of naturally infected milk should be destroyed if the standard conditions are vigilantly maintained.

Malloch & Stoker (1952) found *R. burneti* resistant to treatment with 1% formalin, phenol, Dettol and Cetavlon for 3 hours at various temperatures. Lysol (1%) or ether destroyed nearly all the organisms at 22° C. and 1% Lysol was completely effective after 3 hours at 37° C. *R. burneti* resists the glycerol-phenol treatment used in purification of vaccine lymph from sheep, and animals used for vaccine production at the Lister Institute of Preventive Medicine, Elstree, are now examined routinely for serological evidence of Q fever.

Two strains of *R. burneti* isolated in Great Britain were studied by Stoker (1950) and compared with the classical Henzerling strain isolated from a patient in Italy. The British strains closely resembled the classical strain in morphology, size and behaviour in experimental animals and fertile hen-eggs, and there was reciprocal cross-immunity between the strains in guinea-pigs. Antigens prepared from yolk sacs infected with recently isolated British strains which had been passed only a few times would not fix complement with homologous antisera, despite the fact that the latter reacted well with the Henzerling antigen. After the British strains had undergone several yolk-sac passages, however, they were found to react as well as the Henzerling strain in complement-fixation tests. This change in the behaviour of the strain as a complement-fixing antigen was further investigated (Stoker, 1953) and found to occur suddenly at the fifth yolk-sac passage. Irrespective of the numbers of rickettsiae, antigen made from an earlier passage failed to fix complement with homologous antisera unless the latter was in low dilution. Antigen from later passages, on the other hand, fixed complement with a high serum dilution and in this resembled the Henzerling strain. The early passage antigen, however, was as susceptible to agglutination as the late passage one and both antigens absorbed antibody well, so that the change in behaviour during passage was apparently limited to complement-fixing

ability alone. It is possible that some of the apparent antigenic differences between strains which have been reported (Topping, Shepard & Huebner, 1946) may be due to differences in the stage of adaptation to growth in eggs.

Coombs & Stoker (1951) have shown that the antiglobulin sensitization test is a highly sensitive method for the detection of low concentrations of Q fever antibody. This technique depends on the use of an antiglobulin serum to detect antibody globulin adsorbed on the rickettsial surface. Stoker & Marmion (1952) also used the antiglobulin sensitization test in a study of whey from infected cows. They found that Q fever antibody was present in whey from infected cows, especially from those animals which were excreting rickettsiae, apparently unneutralized, in the milk. The antiglobulin sensitization test was the most sensitive technique for detecting antibody, but of the simpler methods the direct agglutination test on the whey was the most satisfactory and promises to be a convenient method of surveying cattle without having to bleed them or wait for isolation of rickettsiae from the milk.

Little is known of the chemical composition of *R. burneti* but Smith & Stoker (1951) found 9.7% deoxyribonucleic acid and a variable but smaller quantity of ribonucleic acid. The proportions of purines and pyrimidines in the deoxyribonucleic acid resembled those of host (egg) nucleic acid but no 5-methyl cytosine was detectable in the rickettsiae. Smith (personal communication, 1952) and Wyatt & Cohen (1952) have shown that similarly prepared suspensions of *R. mooseri* and *R. prowazeki* have proportions of bases very similar to each other but different from those of *R. burneti* or egg deoxyribonucleic acid. Investigations are now in hand to compare the nucleic acids of different strains of *R. burneti* and the same strain grown in different hosts.

2. Clinical Aspects

Man alone amongst the hosts of *R. burneti* is known to suffer clinically from this infection and the typical illness takes the form of a pneumonia indistinguishable from primary atypical pneumonia of the cold agglutinin type, influenzal pneumonia or psittacosis. Harman (1949), who reviews the clinical findings, points out that suspicion may be aroused by the very sudden onset, severe headache, and absence of upper respiratory and gastrointestinal symptoms, but the diagnosis ultimately rests with the laboratory and usually depends upon demonstration of an increase of specific antibody.

This type of illness accounts for the majority of sporadic cases that would not be diagnosed as Q fever unless they were severe enough to warrant laboratory investigation, but there are probably larger numbers of undiagnosed cases with fever alone and no clinical signs of pulmonary involvement. These fevers, which are usually of short duration, are more often identified as Q fever in outbreaks where all persons at risk are under observation. In fact, they totalled 77% of 43 cases in two outbreaks (Harvey *et al.* 1951; Stoker & Thompson, 1953) but only 29% of 61 sporadic infections (Marmion *et al.* 1953). Admittedly *x*-ray examination might show evidence of pneumonia in the absence of clinical signs, but it is probable that Q fever often causes a mild generalized infection without any localization in the lungs. Increasing experience of the clinical picture of Q fever has also led to the recognition of more severe varieties, especially in older individuals. Four deaths in 69 sporadic cases are reported by Marmion *et al.* (1953) and two (probably three) of these patients had pulmonary infarction. One of them had been febrile intermittently for three months when he died and another had had attacks of fever for several years and had a high titre of Q fever antibody for at least the last six months of his life. These cases, and others reported from abroad, suggest that Q fever may become chronic, and although the disease usually responds to antibiotics such as aureomycin, chloromycetin and terramycin, the severe long-standing cases are often resistant to these drugs.

3. Prevalence of Q Fever in Great Britain

Because the diagnosis depends upon laboratory tests that are available only at a few centres, and because the condition is often mild and susceptible to antibiotics, one would expect that the cases identified as Q fever would represent but a small proportion of the total number which occur. It is difficult to assess the true number and their importance as a cause of morbidity. To judge by specimens received at the Virus Reference Laboratory, Colindale, London, between 1949 and 1951, Q fever accounted for 9% of cases of atypical pneumonia (*Streptococcus* M.G. reactors 16%, influenza 8%, psittacosis 1%) and for 5% of pyrexias of unknown origin (Public Health Laboratory Service, 1953), while only 2 of 110 patients with all types of pneumonia at the Postgraduate Medical School of London had Q fever (Crofton, Fawcett, James & Scadding, 1951). These, however, are selected samples which include a high proportion of patients from urban areas where sources of Q fever are less common. Marmion *et al.* (1953) investigated the rate of infection by determining the proportion of blood donors whose sera contained Q fever antibody, and they found evidence of previous infection in 2.13% of 4,456 donors drawn from England and Wales. In south-east England, where most of the cases have been found, the rate of infection was 2.96% and, after allowance is made for the selection of blood donors by age, it is calculated that this represents some 30,000 infections in Kent alone. These infections might be spread over a long period, because Q fever antibodies persist for many years, but even a conservative estimate would imply a large number of fresh infections in each year. Presumably some of these are asymptomatic as, in outbreaks which have been studied in detail, such cases have totalled up to three times the number clinically ill (Harvey *et al.* 1951).

It is surprising to find that there is no significant difference in the rate of infection between male and female blood donors (Marmion & Stoker, unpublished data, 1952), whereas clinical Q fever is considerably more common amongst males. Even in outbreaks such as that described by Harvey *et al.* (1951), where both sexes were equally at risk, there was a higher attack rate amongst the males. The donor ratios would, therefore, suggest a higher rate of asymptomatic infection among females.

4. Sources and Mode of Infection

Case-to-case infection amongst humans is extremely rare and the only instance of it in Great Britain occurred in a hospital outbreak (Marmion & Stoker, 1950) when two pathologists, a mortuary attendant and a nurse developed Q fever after the death and post-mortem examination of an elderly man with pneumonia. Whittick (1950) found large numbers of rickettsia-like organisms in the lungs of this

patient, but no Q fever occurred amongst the patients or staff who had contact with this patient while he was alive. A similar outbreak has been reported from Italy where 18 doctors and students developed Q fever after a single autopsy (Babudieri, 1951).

Such incidents are rare, and the bulk of human infections are derived from infected domestic animals which remain well but which at parturition excrete large numbers of organisms in the placenta and faeces, and thereafter in the milk. Infection of cows was first found in Kent and Devon by Marmion & Stoker (1950). Slavin (1952) has also detected *R. burneti* in goats' milk in Devon, and Stoker (unpublished data, 1952) has found Q fever antibodies in sheep pastured near infected cattle in Suffolk. Slavin (1952) has carried out a wide survey of cattle in Britain and he has detected *R. burneti* in the milk from 6.9% of 2,581 farms in England, 2.0% of 553 farms in Wales and 0.8% of 240 farms in Scotland.

Marmion *et al.* (1953) found that the rate of infection amongst herds in south-east England was significantly higher than that in East Anglia. This might be linked with the higher rate of infection in blood donors and the larger number of clinical cases in south-east England.

Compared to many other areas, Kent has a large population of sheep, but none of the human cases investigated there so far had obvious contact with these animals or with goats. No history of tick bite has been obtained from patients, but *Haemaphysalis cinnabarina* var. *punctata* is plentiful on the Kent marshes and may play a part in transmission between animals. No strains of *R. burneti* have been isolated from ticks but on two occasions guinea-pigs and mice have developed Q fever antibodies after inoculation with suspensions of *H. cinnabarina*.

Two-thirds of the 69 sporadic cases in Great Britain described by Marmion *et al.* (1953) had had some contact with known sources of Q fever but it is not known what proportion of uninfected individuals are similarly exposed. Of these sporadic cases of Q fever, 32% lived, and 29% worked, near animals or animal products and these cases might have been infected by inhalation of infected dust contaminated with placental or faecal material, while 40% of the patients might have been infected by drinking raw milk. Many drinkers of raw milk live on farms, however, and it is difficult to assess the importance of the milk alone; nevertheless, there were a number of patients in the series whose consumption of infected milk seemed to be the only source of infection and indeed *R. burneti* itself was much more common in the raw milk supplied to these patients than in that supplied to the general public.

Explosive outbreaks of Q fever are presumably due to simultaneous exposure of numbers of susceptible individuals to an unusually large infective dose, and it should be easier to trace the source of these infections than that of isolated cases. When 27 of 185 students at the Canterbury School of Art developed Q fever the most likely source appeared to be dusty packing straw, possibly contaminated with animal excreta, but it was not possible to prove this (Harvey *et al.* 1951).

The main interest of the outbreak on a farm in south Kent lay in the unusually high attack rate (Stoker & Thompson, 1953). Thirteen of fifteen residents developed clinical Q fever, one child had a subclinical infection and one was not tested, and three more cases occurred amongst nine visitors to the farm. On epidemiological grounds, the most probable source was a sow which had littered at the time when infection of the humans must have taken place, but pigs are not a recognized reservoir of Q fever and, possibly owing to the technical difficulty of testing pig sera, it was not possible to prove infection of the animals. After experimental intratracheal inoculation, two pigs developed agglutinins for *R. burneti* but no organisms were detected in their excreta, while a pregnant sow which was similarly inoculated showed a more doubtful serological response and also failed to excrete the rickettsia even after parturition (Stoker & Betts, unpublished data, 1953).

5. Conclusion

Although it is probably common, Q fever is rarely a serious illness and it is usually susceptible to antibiotics, which are now becoming more readily available. Theoretically, it might be possible to eliminate the disease by eradication of infection in domestic animals, but this is so widespread that such measures might be unjustifiably expensive. With more knowledge about the relative roles of milk and air-borne infection, however, as well as about the risks involved in certain occupations, it may be possible to cut down considerably the opportunities for transmission of this infection to man.

References

Babudieri, B. (1951) *Acta med. ital.* **6,** 29

Blanc, G., Martin, L. A., Bruneau, J. & Maurice, A. (1948) *C. R. Acad. Sci., Paris,* **226,** 607

Caminopetros, J. (1948) *Int. Congr. trop. Med.* **4,** 33 [Abstract]

Coombs, R. R. A. & Stoker, M. G. P. (1951) *Lancet,* **2,** 15

Crofton, J. W., Fawcett, J. W., James, D. G. & Scadding, J. G. (1951) *Brit. med. J.* **2,** 1368

Derrick, E. H. (1937) *Med. J. Aust.* **2,** 281

Derrick, E. H. (1944) *J. Hyg., Camb.* **43,** 357

Harman, J. B. (1949) *Lancet,* **2,** 1028

Harvey, M. S., Forbes, G. B. & Marmion, B. P. (1951) *Lancet,* **2,** 1152

Huebner, R. J. & Bell, J. A. (1951) *J. Amer. med. Ass.* **145,** 301

Lennette, E. H. & Clark, W. H. (1951) *J. Amer. med. Ass.* **145,** 306

MacCallum, F. O., Marmion, B. P. & Stoker, M. G. P. (1949) *Lancet,* **2,** 1026

Malloch, R. A. & Stoker, M. G. P. (1952) *J. Hyg., Camb.* **50,** 502

Marmion, B. P. (1951) *J. R. sanit. Inst.* **71,** 97

Marmion, B. P., Rowlands, A., MacCallum, F. O. & Thiel, C. C. 1951) *Mon. Bull. Minist. Hlth Lab. Serv.* **10,** 119

Marmion, B. P. & Stoker, M. G. P. (1950) *Lancet,* **2,** 611

Marmion, B. P., Stoker, M. G. P., McCoy, J. M., Malloch, R. A. & Moore, B. (1953) *Lancet,* **1,** 503

Parker, R. R., Bell, E. J. & Stoenner, H. G. (1949) *J. Amer. vet. med. Ass.* **114,** 124

Public Health Laboratory Service. Virus Reference Laboratory (1953) *Lancet,* **1,** 85

Robbins, F. C. & Ragan, C. A. (1946) *Amer. J. Hyg.* **44,** 6

Slavin, G. (1952) *Vet. Rec.* **64,** 743

Smith, J. D. & Stoker, M. G. P. (1951) *Brit. J. exp. Path.* **32,** 433

Stoker, M. G. P. (1949) *Lancet,* **1,** 178

Stoker, M. G. P. (1950) *Lancet,* **2,** 616

Stoker, M. G. P. (1953) *J. Hyg., Camb.* **51,** 311

Stoker, M. G. P. & Marmion, B. P. (1952) *J. Hyg., Camb.* **50,** 1

Stoker, M. G. P. & Thompson, J. (1953) *Lancet,* **1,** 137

Topping, N. H., Shepard, C. C. & Huebner, R. J. (1946) *Amer. J. Hyg.* **44,** 173

Whittick, J. W. (1950) *Brit. med. J.* **1,** 979

Wyatt, G. R. & Cohen, S. S. (1952) *Nature, Lond.* **170,** 846

THE POST-INFECTION ENCEPHALITIDES

E. WESTON HURST D.Sc. M.D. F.R.C.P.

Biological Laboratories
Imperial Chemical Industries Limited
Manchester

Widespread interest in post-vaccinial encephalomyelitis was first aroused in Great Britain by the classical description of seven fatal cases by Turnbull & McIntosh (1926)[a]; these authors having acceded to an official request to delay publication, their report was forestalled by those of Bastiaanse, Therburgh, Byl & Levaditi (1925)[a] and Lucksch (1925)[a]. Turnbull and McIntosh's first case dated back to 1912; the remainder occurred during 1922 and 1923 and represented the initial impact of a very considerable wave of this type of nervous disease, which caused much alarm in public health circles in the 1920's and early 1930's. (See Ministry of Health, 1928, 1930; Bijl, 1928.)[a]

Turnbull and McIntosh very naturally inquired whether the highly characteristic lesions of post-vaccinial encephalitis had been recognized previously. As stated, one of their own cases had been laid aside as an unusual entity in 1912, and a study of the literature convinced them that a similar condition had occasionally followed smallpox from at least 1874 onwards (Westphal, 1874)[b]. However this may be, it was not long before well-authenticated new cases appeared as a sequel to mild smallpox (McIntosh, 1928; McIntosh & Scarff, 1928; Troup & Hurst, 1930; Brouwer, de Jongh & Rochat, 1933)[b]; Marsden & Hurst (1932)[b] collected a series of 10 further cases and reviewed the older literature in some detail.

Within the same decade an identical disease was recognized as a complication of measles (Bergenfeldt, 1924; Mosse, 1926; Brock, 1926; Ford, 1928; and others)[c]; here again inquiry disclosed a number of more or less well-authenticated cases among earlier records (Barlow & Penrose, 1887; Dreschfeld, 1894; Eliot, 1900; Box, 1921)[c].

At various times, also, the condition has complicated rubella (Bénard, 1921; Davison & Friedfeld, 1938)[d], scarlatina (Winkelman, 1942)[d], varicella (Winnicott & Gibbs, 1926; Glanzmann, 1927; Wilson & Ford, 1927; Zimmerman & Yannet, 1931)[d], and febrile respiratory illnesses (Greenfield, 1930; Davison & Brock, 1937)[d]. In a few instances no antecedent illness has attracted attention (Küssner & Brosin, 1886; Francotte, 1890; Krabbe, 1913; Swan, 1945)[d]. Still other precursors of this type of encephalomyelitis are antirabic treatment (Babes & Mironesco, 1908; Fielder, 1916; Stuart & Krikorian, 1930; Bassoe & Grinker, 1930; and others)[e] and treatment with antitetanus serum (Csermely, 1950)[d]. Finally, rather similar or less closely similar conditions have followed therapy with sulphonamides (Schubert, 1938; Fisher & Gilmour, 1939)[d] and arsenicals (Russell, 1937)[d]; through the haemorrhagic encephalitis considered more characteristic of the neurotoxic effects of arsenic (see Globus & Ginsburg, 1933)[d], these cases possibly link up with the acute haemorrhagic leuco-encephalitis described by Hurst (1941)[d] and with intermediate cases occurring spontaneously.

In the 1920's the incidence of post-vaccinial encephalitis in England was, perhaps, 1 case in 50,000 vaccinations. The disease did not often follow vaccination in infancy or revaccination in later life; nearly all accidents were in those vaccinated for the first time in late childhood or in adult life (Ministry of Health, 1928, 1930)[a]. In Holland, where primary vaccination was ordinarily performed when the child reached school-age, cases were more numerous (Bastiaanse, 1932)[a], whereas some countries insisting on infant vaccination remained relatively free (Ministry of Health, 1928, 1930)[a]. Five of Marsden and Hurst's cases after smallpox occurred among 2,400 patients. The incidence in measles has been given as 1 case in 1,000–1,500 patients (Peterman & Fox, 1933; Geiger & Sappington, 1940; Litvak, Sands & Gibel, 1943)[c]; 75% of cases are in children over the age of five years (Hoyne, 1939)[c]. The frequency of encephalitis is not related to the severity of the preceding exanthem. The " paralytic accidents " of antirabic treatment vary from 1 in 200 to 1 in 8,000 treatments (Koritschoner & Schweinburg, 1925; McCoy, 1930; Horack, 1939; Pait & Pearson, 1949; Latimer, Webster & Gurdjian, 1951; see also Galloway, 1943)[e]; their frequency depends partly on the type of vaccine employed, and is greater among those receiving more intensive treatment (Appelbaum, Greenberg & Nelson, 1953)[e]. The value of all these statistics is relative, however, since very mild cases probably occur and escape diagnosis, and because the disease seems to vary in frequency from decade to decade (Greenfield, 1929[c]; Marsden & Hurst, 1932[b]; Jacobsson & Holmgren, 1949[c]). Moreover, for antirabic treatment the figures include conditions other than that under present consideration (Hurst, 1932)[f]. Occasionally more than one member of a family may suffer, as, for instance, a pair of dissimilar twins vaccinated at the same time (Marsden & Hurst, 1932)[b]. In the USA, post-measles encephalitis is said to affect white people much more commonly than black, and girls more than boys (Hoyne, 1939)[c]. Girls and boys are about equally affected after vaccination (Ministry of Health, 1928, 1930)[a], whereas more males suffer after antirabic treatment (Koritschoner & Schweinburg, 1925)[e]. A second attack is possible and may end fatally (van Bogaert, 1950)[d].

The onset of nervous symptoms, which may be either predominantly encephalitic or myelitic, is abrupt rather than insidious and takes place at a relatively fixed time after vaccination or after the initial manifestations of the exanthem —10 to 13 days after vaccination, 5 to 13 days after the rash of smallpox, at about the end of the first week in measles, during the second half of antirabic treatment, and so on. At this time the patient has developed some immunity to the agent of the primary illness. Paralysis, drowsiness, coma and loss of control of sphincters are among the more common of the many symptoms that may persist for days or several weeks. Temporary blindness is a rare occurrence (Koenigsfeld, 1945[e]; Jennings, 1952[e]). A striking feature of the post-vaccinial cases of the 1920's was the astonishingly rapid and complete recovery of apparently moribund subjects, and the absence of permanent sequelae was contrasted with their distressing frequency in encephalitis lethargica, which had dominated the scene a few years earlier. Clinical recovery is

[a b c d e f] These letters denote the sections of the bibliography in which the reference will be found.

accompanied by ultimate disappearance of the cerebral lesions (Herkenrath, 1935[a]; van Bogaert, 1950[d]). Recent accounts from the USA confirming the impressions of Beach (1895)[c] and Ford (1928)[c] paint a less rosy picture for post-measles encephalitis; many children, variously estimated at from 20% to 90%, show at least a relative impairment of the higher intellectual functions (for example, diminished ability to acquire new knowledge) and of personality and character (Hamilton & Hanna, 1941; Reisman & Rosen, 1943; Appelbaum, Dolgopol & Dolgin, 1949; Sawchuk, La Boccetta, Tornay, Silverstein & Peale, 1949)[c]; the longer the patients are observed and the younger the child, the more noticeable these defects are (Litvak *et al.* 1943; Meyer & Byers, 1952)[c], and their severity is related to the duration of the acute stages of the malady. Kudelka (1932)[a], on the other hand, found no such defects in children recovered from post-vaccinial encephalitis. Mortality in the post-vaccinial cases was usually estimated at around 50% (Ministry of Health, 1928, 1930)[a], in measles at from 6% to 38% (Ford, 1928; Hoyne, 1939; Geiger & Sappington, 1940; and others)[c], and in antirabic treatment at about 20% (Koritschoner & Schweinburg, 1925)[e]. The mortality from post-varicellar encephalitis is apparently very low.

The lesions of post-vaccinial encephalitis were described by Turnbull & McIntosh (1926)[a], Bastiaanse *et al.* (1925)[a], Bouman & Bok (1927)[a], Perdrau (1928)[a], and others; of post-variolar encephalitis by Troup & Hurst (1930)[b] and Marsden & Hurst (1932)[b]; of post-measles encephalitis by Wohlwill (1928)[c], Greenfield (1929)[c], Zimmerman & Yannet (1930)[c], Ferraro & Scheffer (1931)[c] and Malamud (1937)[c]. The characteristic and stereotyped lesion is the perivenular demyelination, often with slight relative sparing of axons, along considerable stretches of the vessels of grey and white matter; the nerve cells are little affected. Similar demyelination occurs beneath the ependyma of the lateral ventricles, beneath the pia mater of brain and cord and along the anterior septum of the cord. The demyelination tends to be minimal in cases of very brief duration. The myelin breaks down much more rapidly than in Wallerian degeneration (Greenfield, 1930)[d], and in cases surviving several weeks or months there is no evidence that the lesions progress after the initial stages (Greenfield, 1930[d]; Malamud, 1939[c])—an important point of distinction from some other demyelinating diseases. The cellular elements present conform with the stage of the morbid process; commonly the perivascular spaces contain many lymphocytes, and the demyelinated areas lymphocytes and highly polymorphic microglia. Except in the very earliest stages, when a proportion of polymorphonuclear leucocytes has been recorded, the increased cells in the cerebrospinal fluid (c.s.f.) are lymphocytes; protein is increased, sugar and chlorides are variable, and the c.s.f. and nervous system bacteriologically sterile.

No conclusive evidence was ever adduced that this type of encephalomyelitis resulted from a direct attack by the virus of the initial exanthem upon the central nervous system. While vaccinia virus was occasionally isolated from the brain or c.s.f. of fatal cases (Turnbull & McIntosh, 1926; Gildemeister, 1929; and others)[a], and measles virus has once been similarly recovered (Shaffer, Rake & Hodes, 1942)[c], deliberate infection of the brain with vaccinia virus evokes quite other lesions (Hurst & Fairbrother, 1930)[f]. An alternative hypothesis that vaccinial infection exposes the nervous system to invasion by a second virus received no support from experiments in which poliomyelitis and vaccinia were injected into the same animal (Hurst & Fairbrother, 1931)[f], nor was a second virus ever isolated to substantiate the view that a latent infection might have been activated by vaccinia. The final hypothesis, and one that has gained favour from recent experimental work, was advanced by Glanzmann (1927)[d] and strongly supported by Finley (1938)[d]. It postulates that the encephalitis is a manifestation of allergy. The experimental work leading to the nearly general acceptance of this view has been reviewed in detail by Hurst (1952)[f] and will be but briefly summarized here.

The fact that antirabic treatment is one of the factors predisposing to this type of encephalomyelitis led Rivers and his colleagues (Rivers, Sprunt & Berry, 1933; Rivers & Schwentker, 1935)[f] to inject suspensions of brain into monkeys. After many months they and Ferraro & Jervis (1940)[f] succeeded in producing an encephalomyelitis accompanied by demyelination. Morgan (1947)[f] and Kabat, Wolf & Bezer (1947)[f] discovered that rather similar but more acute lesions could be induced more speedily and more certainly with inocula, prepared by the technique of Freund & McDermott (1942)[f], known to facilitate the development of immunity to a variety of antigens. Numerous workers have confirmed these observations in a number of animal species and using homologous or heterologous nervous tissue, but not preparations of other tissues. The lesions have at times closely resembled those of post-vaccinial encephalitis, at others acute disseminated encephalomyelitis or acute haemorrhagic leuco-encephalitis. Waksman & Morrison (1951)[f] have shown that the onset of encephalitis coincides closely in time with the development of dermal sensitivity of the type elicited by tuberculin. Olitsky & Tal (1952)[f] and Tal & Olitsky (1952)[f] have recently identified the effective antigens as proteolipides A and B, which belong to a new group of lipoproteins isolated by Folch & Lees (1951)[f]. Sodium salicylate, which inhibits or diminishes various reactions of hypersensitivity, exercises a prophylactic effect against the experimental encephalitis (Good, Campbell & Good, 1949)[f], as does adrenocorticotrophic hormone (Moyer, Jervis, Black, Koprowski & Cox, 1950)[f] and cortisone (Kabat, Wolf & Bezer, 1951)[f]. There is thus circumstantial evidence that the encephalomyelitis is a manifestation of allergy, although the precise mechanism by which the lesions are produced has yet to be demonstrated.

Most of the facts known about the post-infection encephalitides, and particularly the time which elapses after the initial exanthem, can easily be reconciled with the allergic hypothesis; of course, as Hurst (1952)[f] has pointed out, it is not necessary to invoke an antigen of nervous origin in these cases, even though it is possible for an animal to be sensitized with its own nervous tissue (Kabat, Wolf & Bezer, 1949)[f]. Garrison (1952)[e] has recently claimed that a case arising in the course of antirabic therapy made a dramatic recovery as the result of treatment with cortisone. On the other hand, if the basis of the disease is allergy, it is very difficult to understand the local concentration of cases in both space and time to which many writers attest. This is a difficulty for future workers to solve; at the moment there is no convincing rival hypothesis that will account for the occurrence of a stereotyped disease as a sequel to so many different antecedents.

[a b c d e f] These letters denote the sections of the bibliography in which the reference will be found.

SELECT BIBLIOGRAPHY

***a.* Post-vaccinial encephalitis**

Bastiaanse, F. S. van B. (1932) *Rev. neurol.* **39,** 825
Bastiaanse, F. S. van B., Therburgh, —, Byl, — & Levaditi, — (1925) *Bull. Acad. Méd., Paris*, **94,** 815
Bijl, J. P. (1928) *Zbl. ges. Hyg.* **17,** 449
Bouman, L. & Bok, S. T. (1927) *Z. ges. Neurol. Psychiat.* **111,** 495
Gildemeister, E. (1929) *Dtsch. med. Wschr.* **55,** 1372
Herkenrath, B. (1935) *Z. ges. Neurol. Psychiat.* **152,** 293
Kudelka, O. (1932) *Münch. med. Wschr.* **79,** 379
Lucksch, F. (1925) *Med. Klinik*, **21,** 1377
Ministry of Health. Committee on Vaccination (1928) *Vaccination: report of the Committee.* His Majesty's Stationery Office, London
Ministry of Health. Committee on Vaccination (1930) *Vaccination: further report of the Committee.* His Majesty's Stationery Office, London
Perdrau, J. R. (1928) *J. Path. Bact.* **31,** 17
Turnbull, H. M. & McIntosh, J. (1926) *Brit. J. exp. Path.* **7,** 181

***b.* Post-variolar encephalitis**

Brouwer, B., Jongh, C. L. de & Rochat, R. R. (1933) *Dtsch. Z. Nervenheilk.* **131,** 1
McIntosh, J. (1928) *Brit. med. J.* **2,** 334
McIntosh, J. & Scarff, R. W. (1928) *Proc. R. Soc. Med.* **21,** 705
Marsden, J. P. & Hurst, E. W. (1932) *Brain*, **55,** 181
Troup, A. G. & Hurst, E. W. (1930) *Lancet*, **1,** 566
Westphal, C. (1874) *Arch. Psychiat. Nervenkr.* **4,** 335

***c.* Post-measles encephalitis**

Appelbaum, E., Dolgopol, V. B. & Dolgin, J. (1949) *Amer. J. Dis. Child.* **77,** 25
Barlow, T. & Penrose, F. G. (1887) *Med.-chir. Trans.* **70,** 77
Beach, F. (1895) *Brit. med. J.* **2,** 707
Bergenfeldt, E. (1924) *Acta med. scand.* **61,** 281
Box, C. R. (1921) *Lancet*, **1,** 222
Brock, J. (1926) *Z. Kinderheilk.* **40,** 552
Dreschfeld, J. (1894) *Brit. med. J.* **1,** 1174
Eliot, E. F. (1900) *Lancet*, **1,** 533
Ferraro, A. & Scheffer, I. H. (1931) *Arch. Neurol. Psychiat., Chicago*, **25,** 748
Ford, F. R. (1928) *Johns Hopk. Hosp. Bull.* **43,** 140
Geiger, J. C. & Sappington, E. E. (1940) *Arch. Pediat.* **57,** 313
Greenfield, J. G. (1929) *Brain*, **52,** 171
Hamilton, P. M. & Hanna, R. J. (1941) *Amer. J. Dis. Child.* **61,** 483
Hoyne, A. L. (1939) *Illinois med. J.* **76,** 136
Jacobsson, E. & Holmgren, B. (1949) *Ann. paediat.* **173,** 231
Jennings, G. H. (1952) *Lancet*, **2,** 67
Litvak, A. M., Sands, I. J. & Gibel, H. (1943) *Amer. J. Dis. Child.* **65,** 265
Malamud, N. (1937) *Arch. Neurol. Psychiat., Chicago*, **38,** 1025
Malamud, N. (1939) *Arch. Neurol. Psychiat., Chicago*, **41,** 943
Meyer, E. & Byers, R. K. (1952) *Amer. J. Dis. Child.* **84,** 543
Mosse, K. (1926) *Jber. ges. Kinderheilk.* **62,** 272
Peterman, M. G. & Fox, M. J. (1933) *Amer. J. Dis. Child.* **46,** 512
Reisman, H. A. & Rosen, A. S. (1943) *Amer. J. Dis. Child.* **66,** 597
Sawchuk, S., La Boccetta, A. C., Tornay, A., Silverstein, A. & Peale, A. (1949) *Amer. J. Dis. Child.* **78,** 844
Shaffer, M. F., Rake, G. & Hodes, H. L. (1942) *Amer. J. Dis. Child.* **64,** 815
Wohlwill, F. (1928) *Z. ges. Neurol. Psychiat.* **112,** 20
Zimmerman, H. M. & Yannet, H. (1930) *Arch. Neurol. Psychiat., Chicago*, **24,** 1000

***d.* Encephalitis complicating other exanthemata, etc.**

Bénard, R. (1921) *Bull. Soc. méd. Hôp. Paris*, **45,** 1443
Bogaert, L. van (1950) *J. Neuropath.* **9,** 219
Csermely, H. (1950) *Arch. Neurol. Psychiat., Chicago*, **64,** 676
Davison, C. & Brock, S. (1937) *Bull. neurol. Inst. N.Y.* **6,** 504
Davison, C. & Friedfeld, L. (1938) *Amer. J. Dis. Child.* **55,** 496
Finley, K. H. (1938) *Arch. Neurol. Psychiat., Chicago*, **39,** 1047
Fisher, J. H. & Gilmour, J. R. (1939) *Lancet*, **2,** 301
Francotte, X. (1890) *Arch. Neurol., Paris*, **20,** 46
Glanzmann, E. (1927) *Schweiz. med. Wschr.* **8,** 145
Globus, J. H. & Ginsburg, S. W. (1933) *Arch. Neurol. Psychiat., Chicago*, **30,** 1226
Greenfield, J. G. (1930) *J. Path. Bact.* **33,** 453
Hurst, E. W. (1941) *Med. J. Aust.* **2,** 1
Krabbe, K. (1913) *Z. ges. Neurol. Psychiat.* **20,** 108
Küssner, B. & Brosin, F. (1886) *Arch. Psychiat. Nervenkr.* **17,** 239
Russell, D. S. (1937) *J. Path. Bact.* **45,** 357
Schubert, M. (1938) *Derm. Wschr.* **107,** 1361
Swan, C. (1945) *Med. J. Aust.* **2,** 15
Wilson, R. E. & Ford, F. R. (1927) *Johns Hopk. Hosp. Bull.* **40,** 337
Winkelman, N. W. (1942) *J. Neuropath.* **1,** 363
Winnicott, D. W. & Gibbs, N. (1926) *Brit. J. Child. Dis.* **23,** 107
Zimmerman, H. M. & Yannet, H. (1931) *Arch. Neurol. Psychiat., Chicago*, **26,** 322

***e.* Encephalitis in antirabic treatment**

Appelbaum, E., Greenberg, M. & Nelson, J. (1953) *J. Amer. med. Ass.* **151,** 188
Babes, V. & Mironesco, T. (1908) *C. R. Soc. Biol., Paris*, **64,** 964
Bassoe, P. & Grinker, R. R. (1930) *Arch. Neurol. Psychiat., Chicago*, **23,** 1138
Fielder, F. S. (1916) *J. Amer. med. Ass.* **66,** 1769
Galloway, I. A. (1943) *Trop. Dis. Bull.* **40,** 657
Garrison, S. C. (1952) *Amer. J. Med.* **12,** 135
Horack, H. M. (1939) *Amer. J. med. Sci.* **197,** 672
Koenigsfeld, E. G. H. (1945) *J. R. Army med. Cps*, **85,** 254
Koritschoner, R. & Schweinburg, F. (1925) *Z. ImmunForsch.* **42,** 217
Latimer, F. R., Webster, J. E. & Gurdjian, E. S. (1951) *Arch. Neurol. Psychiat., Chicago*, **65,** 16
McCoy, G. W. (1930) *Publ. Hlth Rep., Wash.* **45,** 1888
Pait, C. F. & Pearson, H. E. (1949) *Amer. J. publ. Hlth*, **39,** 875
Stuart, G. & Krikorian, K. S. (1930) *Lancet*, **1,** 1123

***f.* Experimental studies**

Ferraro, A. & Jervis, G. A. (1940) *Arch. Neurol. Psychiat., Chicago*, **43,** 195
Folch, J. & Lees, M. (1951) *J. biol. Chem.* **191,** 807
Freund, J. & McDermott, K. (1942) *Proc. Soc. exp. Biol., N.Y.* **49,** 548
Good, R. A., Campbell, B. & Good, T. A. (1949) *Proc. Soc. exp. Biol., N.Y.* **72,** 341
Hurst, E. W. (1932) *J. Hyg., Camb.* **32,** 33
Hurst, E. W. (1952) *Amer. J. Med.* **12,** 547
Hurst, E. W. & Fairbrother, R. W. (1930) *J. Path. Bact.* **33,** 463
Hurst, E. W. & Fairbrother, R. W. (1931) *Brit. J. exp. Path.* **12,** 17
Kabat, E. A., Wolf, A. & Bezer, A. E. (1947) *J. exp. Med.* **85,** 117
Kabat, E. A., Wolf, A. & Bezer, A. E. (1949) *J. exp. Med.* **89,** 395
Kabat, E. A., Wolf, A. & Bezer, A. E. (1951) *Fed. Proc.* **10,** 412
Morgan, I. M. (1947) *J. exp. Med.* **85,** 131
Moyer, A. W., Jervis, G. A., Black, J., Koprowski, H. & Cox, H. R. (1950) *Proc. Soc. exp. Biol., N.Y.* **75,** 387
Olitsky, P. K. & Tal, C. (1952) *Proc. Soc. exp. Biol., N.Y.* **79,** 50
Rivers, T. M. & Schwentker, F. F. (1935) *J. exp. Med.* **61,** 689
Rivers, T. M., Sprunt, D. H. & Berry, G. P. (1933) *J. exp. Med.* **58,** 39
Tal, C. & Olitsky, P. K. (1952) *Science*, **116,** 420
Waksman, B. H. & Morrison, L. R. (1951) *J. Immunol.* **66,** 421

SOME VETERINARY DISEASES OF MEDICAL INTEREST

A. W. GLEDHILL Ph.D. M.R.C.V.S.

National Institute for Medical Research
Mill Hill, London

1 Foot-and-mouth disease
2 Canine distemper
3 Rift Valley fever
References

The diseases of domestic animals are of medical interest in at least three ways. Firstly, since man's economy is dependent upon domestic animals, their health as well as affecting his well-being directly is an important factor in shaping his social environment. Secondly, research into the nature and control of animal diseases can precede and be a pattern for similar advances in related human diseases. While many researches into the spontaneous diseases of the small laboratory animals have been carried out for this purpose, some diseases of domestic animals investigated as an end in themselves have strikingly influenced the course of medical research. Lastly, some animal diseases are also human diseases and this applies to many virus diseases such as rabies, psittacosis[1], equine encephalomyelitis, louping ill, Rift Valley fever, Q fever[2], etc. Human beings suffering from these diseases are not the primary hosts which ensure the continued survival of the infective agents, because in none of them does infection normally spread directly from man to man and in most the virus particles invading human beings are doomed to extinction. It is a curious fact that no virus diseases are yet known for which the situation is reversed, that is to say, diseases, perhaps of a mild nature, which spread readily from man to man and which sporadically produce severe or fatal infections in animals.

It is proposed in this paper to give an account of three virus diseases of veterinary importance, each of which is of medical interest in one of the above ways: foot-and-mouth disease as a disease of economic importance, canine distemper as having been a landmark in virus research, and Rift Valley fever as a disease affecting man.

1. Foot-and-Mouth Disease

Foot-and-mouth disease is a highly infectious disease of cattle, sheep, pigs and goats. The wild cloven-footed animals are naturally susceptible and several species of rodents are artificially susceptible. Hedgehogs are naturally susceptible by contact with infected cattle and may sometimes be responsible for spreading the disease among domestic animals (McLauchlan & Henderson, 1947). Occasionally human beings contract the disease by ingestion of infective materials or by contact with infected animals. In most outbreaks, the mortality is low (2–3%) in cattle and slightly higher in pigs and sheep. Occasionally, it occurs in a severe form with a mortality rate of 50–70%. The disease is remarkably contagious, especially during the earlier stages in which the virus is circulating in the blood, and immediately afterwards, when vesicles present in the mouth and on the feet rupture and liberate large amounts of virus. Even with transient contact an animal at the highly infective stage regularly infects susceptible animals. However, affected animals rapidly become no longer capable of spreading the disease by contact and the evidence suggests that after recovery few become carriers. Nevertheless persistence of virus on the hoofs for some time after recovery (Gins & Krause, 1923), and the existence of urinary carriers, have both been described (Waldmann, Trautwein & Pyl, 1931). In spite of the comparatively low mortality, it is hardly possible to exaggerate the economic dislocation that would result from allowing the disease to spread freely in countries having a complex, developed economy. This loss would arise from the fact that for several months after infection animals are uneconomic as milk and meat producers. Moreover, immunity resultant upon infection is not very durable nor, indeed, dependable, on account of the antigenically differing types and sub-types of virus.

The disease is due to a filtrable virus estimated by ultrafiltration to be 8–12 mμ. diam. (Galloway & Elford, 1931). More recently it has been separated by the ultracentrifuge into infective, complement-fixing particles of 20 mμ. and non-infective complement-fixing particles of 6.5 mμ. (Agricultural Research Council, 1952). It is resistant, especially in the tissues, to disinfectants, the most practical being 0.5% to 2% sodium hydroxide (Olitsky, Traum & Schoening, 1928). The methods of the meat trade by which meat is frozen protect for long periods the virus in tissues such as lymph nodes, blood in the large vessels, liver, kidneys and rumen. It does not, however, persist so long in the muscles of carcasses in rigor mortis at the time of freezing (Henderson & Brooksby, 1948). This is due to the fact that the pH of muscle, after rigor mortis has set in, falls between the two pH ranges, viz. 6.5–10 and 2.5–3.5, for which the virus is stable (Galloway & Elford, 1935). The recent practice of freezing meat before commencement of rigor mortis enables the virus to survive longer in muscle, although this danger is partly offset by the rapidity of onset of rigor mortis after thawing. The successful demonstration of small amounts of residual virus in such materials is largely a result of recognizing that tests must be carried out in the species to which the virus is adapted; pigs and guinea-pigs are unsuitable for estimating the virus present in bovine tissues. The virus may also survive for long periods after drying at ordinary temperatures. For example, it has been shown to survive on chopped hay at room temperature (56°–66° F.)[3] for at least 15 weeks and on bran for at least 20 weeks (Ministry of Agriculture and Fisheries, 1927).

Although rabbits and rats are susceptible to the virus, the response of these animals is less regular than that of guinea-pigs, in which the typical disease can be produced by inoculation into the balls of the feet (Waldmann & Pape, 1920). Recently Skinner (1951) has shown that the virus produces a fatal disease, probably myositis, in unweaned mice; that these animals are susceptible to intraperitoneal injections of small doses and that the virus attains a very high concentration in their tissues. The use of this laboratory animal has

[1] See Bedson, p. 226 of this Bulletin.
[2] See Stoker, p. 231 of this Bulletin.
[3] t° F. = ⅘ (t − 32)° C.—Ed.

not yet been fully exploited but it nevertheless appears that, in addition to its cheapness and convenience, it provides a very sensitive method of detecting antibody in sera and may prove a useful source of virus for vaccine production. Growth of the virus in tissue culture was first successfully accomplished by Maitland & Maitland (1931), who used embryonic guinea-pig tissue and clotted guinea-pig plasma. Recently, Frenkel (1950) has shown that the virus may be grown on a larger scale on tongues of cattle slaughtered for food.

Three distinct antigenic types of virus have been recognized for more than 25 years: A and O described by Vallée & Carré (1922) and C described by Waldmann & Trautwein (1926). Strains of these types were equally pathogenic for cattle, sheep and pigs and were differentiated by the fact that animals recovered from infection with one type were fully susceptible to infections with the other types. It was soon suspected that differences of lesser degree might exist between strains of the same type. These studies were facilitated by the discovery that the complement-fixation test could distinguish the types of virus (Ciuca, 1929). At first, guinea-pig antigens and sera were used in the test but its usefulness as an aid to quick diagnosis and typing was greatly increased by modifying the test by using bovine antigens (Traub & Möhlmann, 1946). Strains of virus were then encountered which were variants of the classical types (Traub & Rodriguez, 1944; Traub & Möhlmann, 1946). The test has now been put on a more quantitative basis by Brooksby (1948, 1952), giving it increased accuracy and sensitivity. Using his technique, Brooksby, Galloway & Henderson (1948a) showed that while strains from the Mexican outbreak which began in 1946 belonged to type A, they differed from their stock type A and from each other. Confirmation of these results was obtained by using serum neutralization tests in cattle (Brooksby, Galloway & Henderson, 1948b) and by vaccination experiments (Henderson, Galloway & Brooksby, 1948). More recently this team of workers has found three new types of virus, S.A.T. 1, S.A.T. 2 and S.A.T. 3, from outbreaks in Africa (see Agricultural Research Council, 1952).

The methods of control of foot-and-mouth disease are determined by its highly contagious nature and the resistance of the virus outside its host. In countries geographically separated from other countries an attempt is made to prevent the disease from entering them by controlling the importation of materials likely to introduce it. In spite of regulations of this kind, foci of infection may appear and the disease must then be stamped out by slaughtering all susceptible animals in contact with the infection, disposing of their carcasses, enforcing strict quarantine and preventing the movement of animals in the vicinity of the outbreak. Using methods aimed at preventing the importation of potentially infective materials, the USA, Australia and New Zealand have been free from the disease for many years and carry out no research work on it. In Britain, in spite of regulations controlling the importation of infective materials, foci of disease occur and, during periods in which neighbouring countries are suffering an epizootic, many are recorded, especially in the eastern and south-eastern coastal regions. This difference is probably due partly to the fact that Britain imports large amounts of meat and other potentially contaminated materials and partly to its nearness to the European land-mass on which control of the disease presents greater difficulty. Indeed, circumstantial evidence would seem to incriminate migrating birds as passively carrying the virus to Britain from France and the Low Countries (Wilson & Matheson, 1952). The evidence suggests that the virus does not survive long enough on (or in) birds for them to carry it to Britain from Germany and Denmark. In less favourably situated countries having direct land connexion with regions, however distant, where the disease is not controlled, the number of outbreaks is much greater. It then becomes an economic impossibility to control its spread by a policy of slaughter. In this case, the policy adopted is rapidly to build up a wide ring of immunized animals around outbreaks. For this purpose vaccines are employed consisting essentially of formalin-inactivated virus of the same antigenic type as the invading virus. This method of control presupposes a knowledge of the type of virus threatening a region, it tends to leave an enzootic situation rather than to eradicate the disease from a country, and it is costly. Its operation requires a centre for rapid diagnosis and typing of strains by the complement-fixation technique and for preparation of vaccines; the latter necessitates a supply of uniformly susceptible cattle, a necessity itself presenting great difficulty in a country in which many animals have some immunity.

Whereas several methods of preparing formalin-inactivated vaccines are used, an accurate study of the immunizing power of these vaccines against homologous and heterologous strains and of the duration of the immunity which they confer was impossible until a method had been devised for estimating in cattle the amount of virus in infective extracts for preparing vaccines and the amount of virus used in challenging cattle immunized with them. Henderson (1949) described a method capable of giving statistically significant estimations with a reasonable number of cattle. Using this method, Henderson *et al.* (1948) showed that vaccines prepared against their standard A strain, and capable of immunizing cattle against it, failed to immunize satisfactorily against type A strains from the Mexican outbreak which began in 1946. As already noted, Frenkel (1950) devised an economical method of growing the virus in quantity on the excised tongues of cattle, and vaccines prepared with virus grown in this way have now been compared with those prepared from experimentally infected cattle (Agricultural Research Council, 1952). The tests showed conclusively that vaccines prepared by Frenkel's method are comparable with those from infected cattle. Preliminary tests on vaccines prepared from the tissues of infected, unweaned mice have also given encouraging results. These advances give some grounds for hoping that vaccine production will become simpler and cheaper within the next few years.

It is quite obvious that the control of foot-and-mouth disease can be accomplished successfully only by international co-operation and pooling of resources. Vaccine preparation and testing and the preparation of diagnostic guinea-pig sera would be most satisfactorily undertaken at a centre where the disease is never endemic in the local cattle. Information concerning the types of virus present at any time in different parts of the world and the directions of spread could be handled only by a World Foot-and-Mouth Disease Centre, similar to the World Influenza Centre. Unofficially, this role has largely fallen to the Foot-and-Mouth Disease Research Station at Pirbright, Surrey.

2. Canine Distemper

Canine distemper has been recognized for many years as the most serious disease of dogs, especially of young dogs.

The severity and nature of the symptoms displayed vary greatly and it was originally thought that the term canine distemper might include several diseases. Moreover, it was not known whether the diseases of ferrets and of cats that simulated canine distemper all had a common aetiology. When Dunkin and Laidlaw took up the problem in the early 1920's, two divergent views regarding canine distemper were current. The experimental approach had led Carré (1905) to conclude that the disease was due to a filtrable agent that was incapable of being cultivated upon artificial media and was present in the tissues of affected dogs, especially during the earlier stages of the disease. In contrast, Ferry (1910), M'Gowan (1911) and Torrey & Rahe (1913), studying the natural disease during life and at autopsy, came to the conclusion that the bacterium *Haemophilus bronchisepticus* was its primary cause. This organism is commonly isolated from the respiratory tract of dogs in field outbreaks of distemper and occasionally from normal dogs. When introduced into the nose of dogs, it can produce a syndrome including bronchopneumonia, which is not distinguishable from some outbreaks of distemper. Since its pathogenicity decreased with growth on artificial media and since many dogs were refractory to distemper, the irregular transmission of a distemper syndrome by means of this bacterium did not necessarily destroy the claim that it was the cause of distemper. Furthermore, attempts to confirm Carré's work on the filtrability of the causal agent were not notably successful (Kregenow, 1909), and *H. bronchisepticus* was shown by Ferry (1914) sometimes to be filtrable. Dunkin & Laidlaw (1926a), suspecting the ferret to be susceptible to canine distemper and recognizing its advantages over the dog as an experimental animal, established a breeding colony of ferrets and instituted precautions which limited infection to the experimental animals. The disease was thoroughly studied in ferrets, and six strains of canine distemper were established in them. All strains produced one type of disease, and ferrets immunized against any of them resisted challenge with the others. The ferret disease was freely transmissible to dogs, and the canine disease to ferrets. Their evidence indicated that dog distemper was probably a single entity and that the response of the ferret to infection was essentially similar to that of the dog. These workers also used dogs for experimentation (Dunkin & Laidlaw, 1926b). To this end they established a breeding colony and maintained it free from distemper. The experimental disease in dogs was an acute infectious fever characterized by an incubation period of four days, and by coryza and conjunctivitis, a diphasic temperature response, severe gastrointestinal disturbance and rather slight but variable respiratory symptoms. Nervous symptoms due to encephalitis were sometimes observed. In these experiments the infective materials were usually free from all ordinary bacteria, so that complications due to secondary invaders were absent. The mortality was constantly low, contrasting with natural distemper in which it was often high. In none of their experiments was *H. bronchisepticus* isolated, infective materials were usually free of visible or cultivable bacteria and, finally, the disease was transmitted, not without irregularities, by means of the types of bacterial filters used at that time (Laidlaw & Dunkin, 1926). After trying several inactivating agents Laidlaw & Dunkin (1928a) finally used 0.25% formalin to inactivate the virus as Puntoni (1923, 1924) had done, and used infected spleen as the source of virus. Ferrets immunized with vaccine prepared from infected ferret tissue, and dogs immunized with vaccine prepared from infected dog tissue, became sufficiently resistant to withstand challenge with virulent virus 14 days later, which gave them a durable immunity (Laidlaw & Dunkin, 1928a, 1928b). Field trials of this method of vaccination and also the method by which virus and immune serum were inoculated at separate sites gave satisfactory results and the methods have been successfully practised on a commercial scale.

The satisfactory conclusion of Dunkin and Laidlaw's work seemed to give encouragement at a time when there was general despondency and confusion in regard to virus diseases. Apart from their insistence on breeding stocks free from disease and separated from experimental animals, their adaptation of the virus to the ferret as an experimental animal for investigating the disease, although accomplished without difficulty, may have inspired virologists who were working later to persist in efforts, which have now proved so successful, to adapt viruses to grow in laboratory animals and fertile eggs. Certainly, a few years later in the same laboratory, influenza was shown to be due to a virus by successful adaptation to the ferret (Smith, Andrewes & Laidlaw, 1933). In consequence of the irregular results obtained with the bacterial filters then in use, the filtration experiments of Dunkin and Laidlaw were rather uncertain. It was perhaps no accident that Elford (1931), working in the same laboratory, developed within a few years gradocol membranes as a tool for precisely filtering and estimating the size of virus particles.[4]

Later workers, especially in the USA, turned their attention to the control and prevention of distemper in silver foxes and minks. Green (1939) found that, after 54 passes of canine distemper virus in ferrets, its virulence decreased for foxes and dogs (Green & Swale, 1939) to such an extent that it could be used without immune serum as a living vaccine (distemperoid virus) to confer a lasting immunity upon these species (Green & Carlson, 1945). The use of an avirulent living vaccine as an immunizing agent was simpler and cheaper than Laidlaw and Dunkin's methods. Wider experience with it has shown that some risk of distemper is attendant upon its use, although, in this respect, the older methods may not always be perfectly safe (Cabasso, Burkhart & Leaming, 1951; Haig, 1948b). The matter has been carried a step further by Haig (1948a, 1948b) working with Green's distemperoid virus and by Cabasso & Cox (1949) working with virulent distemper virus. These workers have adapted the viruses to grow upon the chorio-allantoic membrane of fertile eggs and noted its decreasing virulence for its natural hosts. After 26 serial passes, the virulence of Cabasso and Cox's egg strain for the ferret, dog and mink was lost, and these animals appeared to be solidly immunized. None of 308 dogs in one field trial and none of 100 dogs in another developed symptoms as a result of vaccination (Cabasso *et al.* 1951; Scanlon & Fisher, 1951). Of 96 of the vaccinated dogs definitely exposed to natural infection, one became infected with distemper.

Partly as a result of failures with the serum and virus method of vaccination advocated by Laidlaw and Dunkin and partly on clinical grounds, it was suspected from 1945 onwards that there existed, especially in Britain, distemper-like diseases of dogs, other than the distemper studied by Laidlaw and Dunkin. Apart from canine hepatitis (Rubarth, 1947), a distemper syndrome ("hard pad") was described which differed from true distemper in the development of hyperkeratosis of the pads, in the frequency of a demyelinating inflammatory encephalitis quite unlike the degeneration of scattered nerve

[4] See Fulton, p. 186 of this Bulletin.

cells described by Laidlaw & Dunkin (1926), and in the susceptibility to it of dogs vaccinated against distemper (McIntyre, Trevan & Montgomerie, 1948). That this disease might be caused by a particularly virulent or somewhat aberrant strain of distemper was suggested by the observation that dogs recovered from natural distemper and dogs which had shown a temperature reaction after vaccination with distemper virus and immune serum were mostly insusceptible to the "hard pad" agent, whereas dogs which showed no temperature reaction after vaccination were susceptible (Greer, 1949). In the USA, Koprowski, Jervis, James, Burkhart & Poppensiek (1950) worked with a strain of distemper virus which also produced a demyelinating encephalitis, but macroscopic evidence of hard pad was absent in their disease. They showed that ferrets vaccinated against distemper with egg-attenuated virus were resistant to their virus. Cabasso (1952) attenuated typical distemper virus, and also the virus of Koprowski and his colleagues, by growing them in eggs, and immunized ferrets with the attenuated variants. Such ferrets were immune to challenge with both heterologous and homologous virulent viruses. Serum against distemper and serum from London against "hard pad" neutralized equally both viruses. Thus, "hard pad" virus differs little from Laidlaw and Dunkin's virus, and animals immunized sufficiently against distemper are insusceptible to "hard pad".

3. Rift Valley Fever

Daubney & Hudson (1931) described an enzootic hepatitis of cattle and sheep in the Rift Valley of Kenya Colony (Rift Valley fever). There was evidence of concomitant infections amongst persons in close contact with affected animals and the paper by these workers includes an account by Garnham of a successful transmission of the disease to a man. Rift Valley fever was shown by them to be due to a small virus which caused a mortality of over 95% in very young lambs, and abortions and a mortality of 20% in ewes and cows. The characteristic lesion was a focal hepatitis, although nephrosis, leucopenia, and some involvement of the mesenteric glands and intestine were also in evidence. The liver cells showed a hyaline degeneration giving a picture closely resembling the Councilman bodies of yellow fever. Intranuclear inclusion bodies characterize the infection (see Rivers, 1952). In man the principal symptoms are headache, backache, anorexia and fever, which last from two to six days (Smithburn, Mahaffy, Haddow, Kitchen & Smith, 1949). No uncomplicated cases are known to have been fatal; nevertheless, defective vision has been described as a not uncommon complication (Freed, 1951; Shrire, 1951). Symptoms referable to liver dysfunction appear to be absent.

Daubney & Hudson (1931) produced strong circumstantial evidence that a vector, probably a mosquito active at night, is responsible for transmitting the disease in sheep and cattle. They did, in fact, produce an inapparent infection in a lamb by injecting an extract of mosquitoes of species *Taeniorhynchus fuscopennatus* caught in the neighbourhood of an outbreak. In 1944 the virus was isolated on six occasions from mosquitoes caught in the Semliki forest in Uganda (Smithburn, Haddow & Gillett, 1948) and it appeared that the virus might have a cycle in the *Eretmapodites* spp. Mosquitoes of the *Eretmapodites chrysogaster* Graham group were then shown experimentally to transmit the disease from mouse to mouse and lamb to lamb (Smithburn, Haddow & Lumsden, 1949). Since the Semliki forest is an uninhabited area, the existence of a sylvan cycle involving this group of mosquitoes and wild animals of unknown species seemed probable. These authors failed to transmit the disease by means of *Taeniorhynchus* spp., although for technical reasons their experiments were inconclusive. In 1951 an epizootic of Rift Valley fever occurred in South Africa (Mundel & Gear, 1951). Whether this was due to the introduction of the disease from central Africa or whether it was an accidental spill-over from some cycle of infection in wild animals has not been determined, but the latter would seem to be more likely (Dick, 1953). Neither in the Rift Valley nor in the area of the outbreak in South Africa are mosquitoes of the *Eretmapodites* spp. believed to be present (Gear, de Meillon, Measroch, Harwin & Davis, 1951). In human beings there is no evidence of person-to-person spread. Laboratory workers handling the virus, and persons in close contact with infected animals and carcasses, frequently become infected (Joubert, Ferguson & Gear, 1951; Gear *et al.* 1951). There is no evidence to suggest that human beings are infected by an insect vector (Dick, 1953) and therefore they presumably play no part in the survival of the virus.

Monkeys and many species of rodents are susceptible to the virus (see Rivers, 1952). Mice are susceptible parenterally and intracerebrally. After sufficient intracerebral passes of the virus in this species its neurotropism is enhanced, whereas its hepatropic affinity is lost. After intraperitoneal inoculation of the neurotropic variant, mice and sheep show no signs of illness, yet are immunized against the hepatropic virus (Smithburn, 1949). Ewes immunized in this way at mating confer sufficient passive immunity upon their offspring to carry them over the age at which the disease is highly fatal. This variant would seem capable of becoming a successful vaccine for use in the field.

REFERENCES

Agricultural Research Council (1952) *Foot-and-mouth disease research: interim report.* Her Majesty's Stationery Office, London
Brooksby, J. B. (1948) *Proc. Soc. exp. Biol., N.Y.* **67,** 254
Brooksby, J. B. (1952) *Rep. agric. Res. Coun.* No. 12
Brooksby, J. B., Galloway, I. A. & Henderson, W. M. (1948a) *Proc. Soc. exp. Biol., N.Y.* **69,** 70
Brooksby, J. B., Galloway, I. A. & Henderson, W. M. (1948b) *Proc. Soc. exp. Biol., N.Y.* **69,** 74
Cabasso, V. J. (1952) *Vet. Med.* **47,** 417
Cabasso, V. J., Burkhart, R. L. & Leaming, J. D. (1951) *Vet. Med.* **46,** 167
Cabasso, V. J. & Cox, H. R. (1949) *Proc. Soc. exp. Biol., N.Y.* **71,** 246
Carré, H. (1905) *C.R. Acad. Sci., Paris,* **140,** 689; 1489
Ciuca, A. (1929) *J. Hyg., Camb.* **28,** 325
Daubney, R. & Hudson, J. R. (1931) *J. Path. Bact.* **34,** 545
Dick, G. W. A. (1953) *Trans. R. Soc. trop. Med. Hyg.* **47,** 13
Dunkin, G. W. & Laidlaw, P. P. (1926a) *J. comp. Path.* **39,** 201
Dunkin, G. W. & Laidlaw, P. P. (1926b) *J. comp. Path.* **39,** 213
Elford, W. J. (1931) *J. Path. Bact.* **34,** 505
Ferry, N. S. (1910) *Amer. vet. Rev.* **37,** 499
Ferry, N. S. (1914) *J. Path. Bact.* **19,** 489
Freed, I. (1951) *S. Afr. med. J.* **25,** 930
Frenkel, H. S. (1950) *Amer. J. vet. Res.* **11,** 371
Galloway, I. A. & Elford, W. J. (1931) *Brit. J. exp. Path.* **12,** 407
Galloway, I. A. & Elford, W. J. (1935) *Brit. J. exp. Path.* **16,** 588
Gear, J., Meillon, B. de, Measroch, V., Harwin, R. & Davis, D. H. S. (1951) *S. Afr. med. J.* **25,** 908
Gins, H. A. & Krause, C. (1923) *Ergebn. allg. Path. path. Anat.* **20** (ii), 805
Green, R. G. (1939) *J. Amer. vet. med. Ass.* **95,** 465

Green, R. G. & Carlson, W. E. (1945) *J. Amer. vet. med. Ass.* **107,** 131
Green, R. G. & Swale, F. S. (1939) *J. Amer. vet. med. Ass.* **95,** 469
Greer, F. G. (1949) *Vet. Rec.* **61,** 163
Haig, D. A. (1948a) *Onderstepoort J. vet. Sci.* **23,** 149
Haig, D. A. (1948b) *J. S. Afr. vet. med. Ass.* **19,** 73
Henderson, W. M. (1949) *Rep. agric. Res. Coun.* No. 8
Henderson, W. M. & Brooksby, J. B. (1948) *J. Hyg., Camb.* **46,** 394
Henderson, W. M., Galloway, I. A. & Brooksby, J. B. (1948) *Proc. Soc. exp. Biol., N.Y.* **69,** 77
Joubert, J. D. S., Ferguson, A. L. & Gear, J. (1951) *S. Afr. med. J.* **25,** 890
Koprowski, H., Jervis, G. A., James, T. R., Burkhart, R. L. & Poppensiek, G. C. (1950) *Amer. J. Hyg.* **51,** 63
Kregenow, C. (1909) *Zbl. Bakt.* (1 Abt., Orig.) **50,** 326
Laidlaw, P. P. & Dunkin, G. W. (1926) *J. comp. Path.* **39,** 222
Laidlaw, P. P. & Dunkin, G. W. (1928a) *J. comp. Path.* **41,** 1
Laidlaw, P. P. & Dunkin, G. W. (1928b) *J. comp. Path.* **41,** 209
M'Gowan, J. P. (1911) *J. Path. Bact.* **15,** 372
McIntyre, A. B., Trevan, D. J. & Montgomerie, R. F. (1948) *Vet. Rec.* **60,** 635
McLauchlan, J. D. & Henderson, W. M. (1947) *J. Hyg., Camb.* **45,** 474
Maitland, M. C. & Maitland, H. B. (1931) *J. comp. Path.* **44,** 106
Ministry of Agriculture and Fisheries (1927) *Second progress report of the Foot-and-Mouth Disease Research Committee.* His Majesty's Stationery Office, London
Mundel, B. & Gear, J. (1951) *S. Afr. med. J.* **25,** 797
Olitsky, P. K., Traum, J. & Shoening, H. W. (1928) *Tech. Bu.. U.S. Dep. Agric.* No. 76, p. 93
Puntoni, V. (1923) *Ann. Igiene (sper.)* **33,** 553
Puntoni, V. (1924) *Ann. Igiene (sper.)* **34,** 406
Rivers, T. M., ed. (1952) *Viral and rickettsial infections of man,* 2nd ed. Lippincott, London
Rubarth, S. (1947) *Acta path. microbiol. scand.* Suppl. No. 69
Scanlon, E. J. & Fisher, G. W. (1951) *Vet. Med.* **46,** 432
Shrire, L. (1951) *S. Afr. med. J.* **25,** 926
Skinner, H. H. (1951) *Proc. R. Soc. Med.* **44,** 1041
Smith, W., Andrewes, C. H. & Laidlaw, P. P. (1933) *Lancet,* **2,** 66
Smithburn, K. C. (1949) *Brit. J. exp. Path.* **30,** 1
Smithburn, K. C., Haddow, A. J. & Gillett, J. D. (1948) *Brit. J. exp. Path.* **29,** 107
Smithburn, K. C., Haddow, A. J. & Lumsden, W. H. R. (1949) *Brit. J. exp. Path.* **30,** 35
Smithburn, K. C., Mahaffy, A. F., Haddow, A. J., Kitchen, S. F. & Smith, J. F. (1949) *J. Immunol.* **62,** 213
Torrey, J. C. & Rahe, A. H. (1913) *J. med. Res.* **27,** 291
Traub, E. & Möhlmann, H. (1946) *Berl. Münch. tierärztl. Wschr.* **1,** 1
Traub, E. & Rodriguez, R. (1944) *Zbl. Bakt.* (1 Abt., Orig.) **151,** 380
Vallée, H. & Carré, H. (1922) *C.R. Acad. Sci., Paris,* **174,** 1498
Waldmann, O. & Pape, J. (1920) *Berl. tierärztl. Wschr.* **36,** 519
Waldmann, O. & Trautwein, K. (1926) *Berl. tierärztl. Wschr.* **42,** 569
Waldmann, O., Trautwein, K. & Pyl, G. (1931) *Zbl. Bakt.* (1 Abt., Orig.) **121,** 19
Wilson, W. W. & Matheson, R. C. 1952) *Vet. Rec.* **64,** 541

* * *

Book Reviews

The Nature of Virus Multiplication : Second Symposium of the Society for General Microbiology held at Oxford University, April 1952

Society for General Microbiology. Edited by *Paul Fildes & W. E. van Heyningen.* Cambridge: University Press, 1953. x + 320 pages; plates; figures. 25 × 16 cm. £1 15s.

Interest in the mechanism of virus multiplication has been greatly stimulated in recent years by a wealth of research on animal and plant viruses and on bacteriophage, much of which has a bearing on a problem not only of the greatest intrinsic interest but closely related to the possible development of chemotherapy against infections by these agents.

This book presents the 16 papers which were " taken as read " at the meeting in 1952. These take the form of descriptions of original work, reviews of other findings and attempts to interpret and correlate the results. Some of the contributions are well illustrated by photographs and many of the experiments are shown in figures. After a list of references, the discussions which followed most of the papers are given in an amended form. In all, the views of 45 participants are recorded.

Apart from the diversity of viruses considered, the types of investigation which may throw light on problems of their multiplication vary enormously. Thus infectivity and serological studies of influenza virus, protein and nucleic acid synthesis in plant viruses, mutation and recombination in phages and the morphology of insect viruses and those of the influenza group are some of the aspects to which attention is directed.

In spite of this, much more factual information is required and there is, therefore, a great deal of speculation which, as one critic observes, may at times verge on the metaphysical. Very generally, the concept of multiplication by preliminary increase of sub-units or building blocks is supported by some workers who have made a special study of bacteriophage or influenza virus. This idea stems mainly from the failure to detect infective virus in cells for some time after infection—the so-called dark period—, from the premature increase in soluble antigen and from the phenomenon of genetic recombination. It is necessary to appreciate that these findings can be interpreted differently. In fact this concept of the behaviour of influenza virus has recently been questioned seriously by the Henles who had earlier supported it, and further experiments by Dulbecco suggest that multiplicity reactivation in bacteriophage cannot be explained in this way. Arguments for the persistent integrity of the virus particle and multiplication by fission are considered by various contributors and especially in one paper devoted to this theme. Microscopy of some of the larger viruses provides evidence for this older hypothesis and, in the absence of evidence demanding acceptance of some quite aberrant mechanism for at any rate part of the virus group, it would seem more logical to favour the conservative view-point which is, at least, not disproved and has the advantage of uniformity.

As a whole this publication presents a most interesting collection of experimental results of great diversity. The fact that many different opinions as to their interpretation are put forward only emphasizes the stimulating quality of the book and the excellent arrangement of the Symposium which it records.

C. F. Barwell

Viral and Rickettsial Infections of Man

Edited by *Thomas M. Rivers.* 2nd ed. London: J. B. Lippincott Co., 1952. xvi + 719 pages; 7 plates; 90 illustrations. 26 × 18 cm. £3

There are a number of textbooks in English devoted to virus diseases. The most encyclopaedic is *Virus diseases of man*, by van Rooyen & Rhodes (2nd ed., Nelson, 1948). This work is valuable for its bibliographies, but it can be recommended only to those who already possess a detailed knowledge of the subject. The writers have chosen to present the maze of facts with so little bias that the beginner would not be able to extract the good from the bad. It seems to the reviewer that textbooks should present the current point of view of their author; for even in the most massive textbook it is necessary to select the references from a much wider field, and if a bibliographical source is required it is better to rely on complete lists such as are available for poliomyelitis (*A bibliography of infantile paralysis, 1789–1949: with selected abstracts and annotations*, edited by M. Fishbein & E. M. Salmonsen with L. Hektoen, 2nd ed., Lippincott, 1951). Unfortunately such complete lists do not yet exist for the majority of viruses.

A very short didactic treatment by the same authors (Rhodes & van Rooyen: *Textbook of virology for students and practitioners of medicine*, Nelson, 1949) may be highly recommended to those who wish to devote only a limited time to the subject. In the same category there is an excellent introduction by Patrick Meenan (*The essentials of virus diseases*, Churchill, 1951), though in this book there are no references to the original papers.

Between these two extremes of the encyclopaedic and the succinct lie two textbooks of relatively moderate length. The British textbook is the work of S. P. Bedson, A. W. Downie, F. O. MacCallum & C. H. Stuart-Harris (*Virus and rickettsial diseases of man*, Arnold, 1951); its American counterpart is the volume edited by T. M. Rivers. This work has now reached its 2nd edition and is the main subject of this review. It is more than double the size of the British book, but both deal with the subject from very much the same standpoint.

To those who know the 1st edition of the American textbook it is sufficient recommendation to say that it has now been improved. The individual chapters are written by American virologists well known for their contributions to the subject with which they deal. The book is on an ambitious scale, and although probably too long for undergraduates, it should appeal to postgraduates in the clinical and public health fields and to bacteriologists. In general, viruses are considered as the cause of disease in man, but some of the chapters are devoted to the more academic aspects of the subject: for example, there is a most lucid exposition of the modern knowledge about bacteriophages. With the extensive though critical bibliographies which are appended to each chapter, it would be easy to plunge into the original literature confident of wise guidance.

Most of the articles which appeared in the 1st edition show evidence of revision, though the extent of this naturally varies with the interest focused on the subject in the past few years. This revision and the inclusion of articles by four new contributors have increased the length of the book by over 100 pages. Hirst has written a short article on haemagglutination; Schlesinger discusses interference between animal viruses, and Melnick & Curnen have provided an excellent review of the Coxsackie group of viruses which was not even discovered when the 1st edition appeared. Descriptions of swineherd's disease and of pretibial fever are deleted as they are now known to be caused by *Leptospira*. Thanks to a grant from the National Foundation for Infantile Paralysis, it has been possible to publish this magnificent textbook at a relatively modest price. The printing and reproduction of the many photographs are excellent; in this edition there is a distinct improvement in the quality of the electron micrographs. It can be recommended without reservation.

Forrest Fulton

Expert Committee on Hepatitis : First Report

World Health Organization. Geneva: World Health Organization, 1953. (World Health Organization Technical Report Series No. 62.) 26 pages. 24 × 16 cm. 1s. 6d.

This brief report by a representative and experienced committee does not pretend to be a critical survey of the literature, but deals succinctly with the present state of our knowledge of the hepatitis virus. In brief paragraphs it deals with, among other things, the resistance of the virus to the physical and chemical agents, natural and artificial modes of spread of the virus, specific and general measures of control, prevention of spread of the virus by human blood and its products, and measures to render parenteral injection safe.

For the research worker, the brief reference to viruses causing hepatitis in animals is a most timely warning of the pitfalls waiting for the enthusiastic, as well as a reminder of the important observations of Gledhill and Andrewes on mice and of Rubarth on

dogs. These observations, as the committee rightly point out in their discussion of further fields for research, may well be a source from which fresh and valuable light might be thrown on the problem of the virus infection in humans.

In such a vexatious field it is perhaps unkind to criticize such an excellently lucid, though limited, report. The section on terminology is welcome, but might have been bolder. The term "serum hepatitis" is clearly misleading, and "catarrhal jaundice" certainly wrong. All terms have their limitations. Might it not be simpler in the present state of our knowledge to refer to the two conditions —if two there be—as virus hepatitis A and virus hepatitis B?

Table I, which compares certain features of the two conditions, implies to the uninformed reader a clearer-cut division than experience suggests or the text in this monograph admits. This table could have been omitted. As shown in Table III, the effect on virus A of chlorination, as applied in the field, surely remains equivocal. The observation bleakly recorded in Table II, that there is no evidence that γ-globulin has a prophylactic value against virus B, might have been emphasized. If there is any real substance in the claim that 1 in 200 blood donors carry virus B in their blood, this is a matter of importance. It might have been helpful at the same time to give tentative figures indicating the risk of jaundice through whole blood transfusion, to put the matter in perspective.

It would be unjust to end without praising the committee for their excellent report in a most difficult field.

Nicholas H. Martin

Poliomyelitis

W. Ritchie Russell. London: Edward Arnold & Co., 1952. vii + 84 pages; 20 figures. 22 × 14 cm. 14s.

In this admirable monograph Dr Ritchie Russell covers most aspects of poliomyelitis and endeavours to present the disease as a whole. A surprising amount of information is contained in less than 100 pages. The larger part of the book is devoted to the diagnosis and treatment of the disease; as a result, certain aspects, such as the method of spread of infection and the promising field of immunization, are dealt with very briefly.

The chapters on the clinical features of the disease and the physical examination of patients are particularly well done. The author presents a series of case histories drawn from his own wide experience, which clearly illustrate the varied symptoms and signs. The necessity for early diagnosis is emphasized. Dr Ritchie Russell points out that at one time this was of limited practical importance; now, however, with the development of specific treatment for several acute intracranial infections (e.g., tuberculous meningitis), early and accurate diagnosis has become increasingly important. The importance of absolute mental clarity and the appearance of muscular fibrillation in the pre-paralytic stage of poliomyelitis is emphasized.

Dr Ritchie Russell's pioneer work on the importance of complete rest in the acute phase of the disease is now generally accepted. His further suggestion that school sports and other severe forms of physical exercise should be discouraged during the poliomyelitis season, although logical, might well cause undue public anxiety. The author rightly condemns panic measures such as the removal to hospital of all possible cases of poliomyelitis; such a physical and mental disturbance might well turn a non-paralytic illness into a paralytic one, particularly in children.

The prime importance of maintaining a clear airway is stressed; with the skilled use of postural drainage, tracheotomy should rarely be necessary. Various types of respirator are described. The author clearly favours the tank-type machine and suggests many ways in which it can be improved.

Most authorities will agree with the author's views that gentle passive movements should be started as soon as possible. It is questionable whether his subsequent advice that muscles should, within a few weeks, be exercised to the point of complete fatigue will be acceptable to the majority of his orthopaedic colleagues. The "treat-the-muscle-gently" school are unlikely to be influenced by the evidence produced by Dr Ritchie Russell in favour of early strenuous active exercise.

The writer draws attention to many unsolved problems and makes constructive suggestions as to how these problems can be tackled. He strongly advocates special centres for the treatment and study of patients with poliomyelitis. Experience has already shown how inadequate are the existing arrangements in Britain—a country which up to now has been largely spared the ravages of this cruel disease.

The book is clearly printed and the photographs are good. Dr Ritchie Russell, in covering so many aspects of the subject in a small book, successfully fills a gap in the vast literature of poliomyelitis.

W. P. Sweetnam

Poliomyelitis: Papers and Discussions Presented at the Second International Poliomyelitis Conference

International Poliomyelitis Congress. London: J. B. Lippincott Co., 1952. xxi + 555 pages; illustrations. 26 × 18 cm. £3

This volume reports the proceedings of the Second International Poliomyelitis Conference, held in September 1951 in Copenhagen, and as such forms an invaluable work of reference to our knowledge of poliomyelitis at that time. The book should be in the hands of all those concerned in the study of this disease.

The volume opens with stimulating addresses by Professor Niels Bohr, president of the Conference, and Mr Basil O'Connor, president of the International Poliomyelitis Congress and of the National Foundation for Infantile Paralysis, who has done so much to further research on poliomyelitis and dissemination of knowledge regarding it.

The first symposium is devoted to "Virus and its interaction with the host cell". C. H. Andrewes writes on "Viruses as organisms", and W. M. Stanley presents the arguments in favour of regarding viruses as chemical agents. Other contributors include F. C. Bawden, Max Delbrück and F. L. Horsfall.

The pathology and pathophysiology of poliomyelitis are dealt with in the first review section. The first paper, the most important presented at the Conference, is written by J. F. Enders, and describes the cultivation of poliomyelitis virus in tissue cultures of human cells. Reference is made to virus isolation, typing, and the possibilities of using this technique in the search for a virus vaccine. This paper points to the necessity for revising the concept that poliomyelitis virus is a strict neurotrope. The next paper, by D. P. C. Lloyd, on neurophysiology, discusses responses of motoneurones. David Bodian writes an authoritative article, "Virus and host factors determining the nature and severity of lesions and of clinical manifestations". This article is illustrated by many excellent photographs and charts. Ruth E. M. Bowden writes on studies of skeletal muscle in poliomyelitis, and P. Bauwens discusses electromyography and the motor unit.

The second symposium is devoted to two very complete papers on Coxsackie infections, by G. Dalldorf and E. C. Curnen. These papers describe the characteristics of the viruses and present the evidence that they cause human disease. Then follow the reports of the official delegates, who outline the epidemic pattern in their own countries.

The third symposium is concerned with problems of differential diagnosis. R. Debré & S. Thieffry discuss neurological aspects, with particular reference to paralytic disease, and J. E. Smadel & C. V. Adair laboratory aspects, with main reference to non-paralytic poliomyelitis. J. Casals then describes his work, quite new at that time, on complement fixation. Finally J. E. Salk presents a report on the typing of 100 strains of poliomyelitis virus into three main types.

The second review section is on the treatment of poliomyelitis: J. L. Wilson writes on the management of respiratory insufficiency, H. J. Seddon on the reconstructive surgery of the upper limb, C. E. Irwin on the surgical rehabilitation of the hand, J. R. Cobb on the correction of scoliosis, R. L. Bennett on physical medicine, A. Bruun on social problems, and M. A. Seidenfeld on psychological problems.

The fourth symposium opens with a discussion on immunity. H. A. Howe writes on the role of antibodies, and J. L. Melnick on latent survival of viruses. W. R. Russell describes how the course of the illness can be influenced by physical activity, and A. B. Hill writes on inoculation procedures as provoking factors.

The third review section is devoted to ecology and epidemiology. J. H. S. Gear discusses the possibilities of extrahuman survival of the poliomyelitis virus, T. Francis writes on the distribution of the virus in infected communities, and G. Olin describes the interesting epidemiological pattern in Sweden.

Finally, J. R. Paul writes a summary on "Knowledge and trends in poliomyelitis". He emphasizes the extreme value of Enders' discovery that poliomyelitis virus grows in tissue culture. This appraisal has been amply borne out in the two years since the

Conference was held. Paul suggests that the end of " the monkey era " is at hand. While this may be true as regards experimentation on pathogenesis and isolation tests, this animal is still invaluable as a source of tissue for tissue cultures. Paul looks forward to the diagnostic possibilities opened up by the development of a complement-fixation test. It has subsequently become clear that this test is not likely to be of much value in diagnosis, as responses are broad, and antibody reaches high titre early in the illness. Attention is then drawn to the important work on typing.

In closing, Paul refers to the changing history of poliomyelitis. No longer is the disease regarded as one of infants. Originally believed to be mildly contagious, poliomyelitis is now held to be almost as contagious as measles. The clinical picture of paralysis is now thought to be the rarest, not the commonest, manifestation of infection. Paul concludes with these words: " We have an obvious and great cause and if there can be developed . . . new technics . . . during the next three years . . . we may have good reason to be optimistic for the future ". Recent reports on the value of tissue cultures in the isolation and identification of virus, on the use of γ-globulin as a prophylactic, and on the development of a poliomyelitis vaccine, justify the optimism expressed by Paul and felt by most of those who were privileged to attend the Conference.

A. J. Rhodes

Influenza and Other Virus Infections of the Respiratory Tract

C. H. Stuart-Harris. London: Edward Arnold & Co., 1953. vi + 235 pages; 98 figures. 22 × 14 cm. £1 10s.

For many years virus diseases of the respiratory tract were customarily diagnosed as influenza. This term, being indiscriminately used, in fact included a number of different diseases of varying severity, the actual diagnosis of each being regarded as relatively unimportant, since the treatment was purely symptomatic. The advent of antibiotics has fundamentally changed this attitude, and it is fortunate that this book by Professor Stuart-Harris has appeared. He is well qualified for this purpose since, in addition to being an expert in clinical medicine, he has a profound knowledge of the complicated laboratory procedures necessary for the identification of the different infecting agents.

This book brings together in a masterly fashion the whole of the present state of knowledge. There are chapters on the clinical picture of influenza, the chest complications (illustrated with first-class radiographs), pathology, laboratory diagnosis, differential diagnosis, prevention and treatment, whilst the chapter on epidemiology is most fascinating. Information is included about psittacosis, ornithosis, Q fever and the common cold; virus pneumonia has a chapter to itself, again illustrated with excellent radiographs.

The author rightly discourages the indiscriminate use of antibiotics, so prevalent at the present time, and stresses that their use should be restricted to serious cases, in which adequate dosage of the correct material must be employed.

It is often remarked that there is no interest left in the work of a general practitioner—merely the neurotics and chronics remain. To read the chapter on epidemiology is to realize how much of our present knowledge of the spread of influenza we owe to an observant practitioner in the dales of north Yorkshire, whereas at the same time the importance of international co-operation in health matters is stressed. It is interesting to read that the infecting agent of the 1918–19 influenza pandemic is still not identified. The wide scale on which investigations of the whole subject have been conducted is described; this is understandable, considering the colossal man-power wastage for which this group of diseases is responsible.

This book is well produced and can be thoroughly recommended to all those interested in the accurate treatment and diagnosis of these diseases. In conclusion, one may hope that, with new advances still to be made in this important subject, further editions of this book may become available under the capable hand of the author.

W. Smith

Lecture Notes on the Use of the Microscope

R. Barer. Oxford: Blackwell Scientific Publications, 1953. vii + 76 pages; 8 figures. 19 × 12 cm. 6s.

This book is a brief and simple explanation of how an ordinary laboratory microscope works. It contains clear and simple directions on how to set the light, adjust the condenser, select the appropriate combination of objective and eyepiece, and focus: in short, how to get the best out of a microscope.

Students, laboratory technicians and many graduates will find it useful. The chapter on the cleaning of microscope lenses is likely to be specially helpful to those in charge of large classes. A simple account is given of numerical aperture, and of the role of diffraction in image formation, with experiments designed to illustrate their importance; and of dark-ground illumination by means of a central stop in an ordinary condenser. It is a pity that no instruction is given on the use of the oil immersion dark-ground condenser. Many graduates as well as students find this difficult.

The book ends with a description of Köhler illumination and the use of the " research " type of microscope lamp. It may be hoped that future editions will contain a simple explanation of the phase contrast microscope. Students often see objects demonstrated in this way even if they may not use the phase microscope themselves, and are apt to ask questions—awkward if the demonstrator doesn't know the answers.

For six shillings this book is a bargain.

T. H. Flewett

Lectures on the Scientific Basis of Medicine. Volume I: 1951–52

British Postgraduate Medical Federation, University of London. London: Athlone Press, 1953. vii + 396 pages; 6 plates; figures. 22 × 14 cm. £1 10s.

The publication of these selected lectures, given before the British Postgraduate Medical Federation during the winter of 1951–52, will be of great interest not only to the medical research scientist, but also to the open-minded clinician who finds himself unable to keep track with the wealth of information published each year in the various spheres of medical research. The lectures deal with recent advances in medical science and their authors are well known for their scientific achievements.

In the introductory lecture, Professor Adrian discusses the scientific approach to medical research. He warns the medical scientist of the difficulties he will have to overcome. Man is not an easy animal to experiment on:

> . . . our biological material is the most difficult of all. It is not merely alive but it is as conscious and intelligent as we are. And since our material is man we cannot quite avoid approaching it with a slightly emotional bias . . .

Though the difficulties are very great, tremendous advances have been made by the application of scientific methods to medicine.

Of special interest to the microbiologist and to the virologist are the lectures by Dr Miles, Professor Bedson and Professor Downie.

Dr Miles' lecture deals with some aspects of antibacterial immunity, a field to which he has made very important contributions. In a language where much of the technical jargon has been avoided he discusses: the simple vaccine; the search for immunizing antigens; the toxins; criteria of toxicity of bacteria; the endotoxins; the auxiliary pathogenic factors; immunity to *Str. pyogenes*, *Staph. pyogenes* and *B. anthracis*; passive immunity.

Professor Bedson summarizes very clearly and in simple words the latest theories on viruses as the causes of disease. The modes of spread of virus infection, the spread of infection within the host, the changes produced in the cells and tissues by viruses and the effect on the host of the cellular changes produced by viruses are concisely exposed and explained by well-selected examples.

Professor Downie tells us about antibodies and immunity to virus infection. The knowledge acquired in the past quarter of a century on the immunology of virus diseases has led to the development of valuable diagnostic techniques and to the application of important preventive measures. Professor Downie's comprehensive review covers such topics as: resistance to virus infection independent of specific antibody; the prevention of infection by immune serum; correlation of serum antibody levels and resistance to infection; the estimation of antibody; mode of action of antibody; the development of antibody and the course of infection; antibody duration and immunity; subclinical infection.

As the present review is intended for the Viruses in Medicine number of the *British Medical Bulletin*, emphasis has been put more particularly on the lectures dealing with viruses. However, it must be noted that the other lectures, covering various fields of physiology and physiopathology, will be of great interest not only to the specialist but also to the non-initiated.

Paul Fiset

Printed in Great Britain by Spottiswoode, Ballantyne & Co. Ltd., London & Colchester

DOES NOT CIRCULATE

Medical

OCT 16 1953

BRITISH MEDICAL BULLETIN

Viruses in Medicine

INTRODUCTION	*C. H. Andrewes*
ECOLOGY OF VIRUS DISEASES	*Sir Macfarlane Burnet*
PROGRESS IN VIRAL IMMUNOLOGY	*Wilson Smith*
CHEMOTHERAPY OF VIRUS DISEASES	*E. Weston Hurst*
TECHNIQUES FOR THE STUDY OF VIRUSES	*Forrest Fulton*
SMALLPOX AND RELATED VIRUS INFECTIONS IN MAN	*A. W. Downie & A. Macdonald*
RECENT ADVANCES IN THE STUDY OF POLIOMYELITIS	*A. J. Rhodes*
COXSACKIE VIRUSES	*J. O'H. Tobin*
THE COMMON COLD	*C. H. Andrewes*
THE VIRUSES OF EPIDEMIC INFLUENZA	*A. Isaacs*
MUMPS	*E. T. C. Spooner*
YELLOW FEVER	*G. W. A. Dick*
HEPATITIS	*F. O. MacCallum*
PSITTACOSIS-LYMPHOGRANULOMA VIRUSES	*S. P. Bedson*
Q FEVER IN BRITAIN	*M. G. P. Stoker*
THE POST-INFECTION ENCEPHALITIDES	*E. Weston Hurst*
VETERINARY DISEASES OF MEDICAL INTEREST	*A. W. Gledhill*

★ ★ ★

BOOK REVIEWS

Published by Medical Department, The British Council, an agent of the British Government

VOLUME 9 NUMBER 3 1953
(Completing the Volume)

PRICE $2.75

Increased Life-span for Virus Suspensions . . . Vaccines Cultures, etc., WITHOUT External Refrigeration

OTHER VACUUM EQUIPMENT AVAILABLE FOR LABORATORY AND HOSPITAL

"SPEEDIVAC" Rotary Vacuum Pumps.

●

Complete Mobile Vacuum Systems.

●

Shadow-casting equipment for microscopists.

●

Tissue Dryers for the instantaneous 'Fixation' and dehydration of plant and animal tissue.

●

Laboratory Vacuum Ovens.

FREEZE DRYING—a technique for which Edwards have developed a complete range of equipment for medical and veterinary laboratory—has made possible many wonderful advances in medicine and surgery.

The life of Vaccines, Serum, Plasma, Cultures and Tissues can be suspended at will, their span of activity increased indefinitely and their stability improved without the need for complicated refrigerant equipment—a valuable factor where long distance transportation or prolonged research work is involved.

Precision equipment is also available for the micro-moisture determination of biological materials and pharmaceuticals.

PLEASE WRITE FOR FURTHER INFORMATION

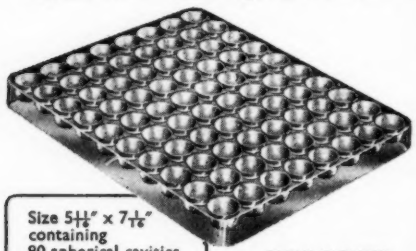
for better vacuum service . . .

W. EDWARDS
& CO. (LONDON) LTD., LONDON, S.E.26
Tel.: Sydenham 7026 (8 lines). Grams.: Edcohivac, Souphone, London

As used by World Influenza Centre

A PLASTIC PLATE FOR USE IN TESTS INVOLVING VIRUS HÆMAGGLUTINATION & OTHER SIMILAR REACTIONS

Supplied to clinics throughout the world



Size 5 11/16" x 7 1/16" containing 80 spherical cavities 3/4" diameter

EASILY CLEANED MOULDED CUPS GLASS SMOOTH

SOLE MANUFACTURERS:—

Prestware Ltd.
(B DIVISION)
LOMBARD ROAD, MERTON, LONDON, S.W.19
Tel: LIBerty 3421

E. K. BOWMAN
LTD

Established 1889

61, 61A & 63 HOLMES ROAD,
KENTISH TOWN, LONDON, N.W. 5

Telephone: Gulliver 1653

★

SPECIALISTS IN ALL TYPES OF
ANIMAL CAGES
TUBULAR STEEL AND ANGLE IRON RACKING
CAGES FOR MONKEYS, BABOONS AND THE SMALL RODENT

★

SEND FOR ILLUSTRATED CATALOGUE